DOCUMENT RESUME

ED 255 933 CS 208 842

AUTHOR	Fagan, William T.; And Others
TITLE	Measures for Research and Evaluation in the English Language Arts. Volume 2.
INSTITUTION	ERIC Clearinghouse on Reading and Communication Skills, Urbana, Ill.; National Council of Teachers of English, Urbana, Ill.
SPONS AGENCY	National Inst. of Education (ED), Washington, DC.
REPORT NO	ISBN-0-8141-3101-8
PUB DATE	85
CONTRACT	400-83-0025
NOTE	248p.; For the first volume, see ED 099 835.
AVAILABLE FROM	National Council of Teachers of English, 1111 Kenyon Rd., Urbana, IL 61801 (Stock No. 31018, $13.00 member, $16.75 nonmember).
PUB TYPE	Tests/Evaluation Instruments (160) -- Information Analyses - ERIC Information Analysis Products (071)
EDRS PRICE	MF01/PC10 Plus Postage.
DESCRIPTORS	Attitude Measures; *Educational Assessment; *Educational Research; *English Instruction; Language Acquisition; *Language Arts; Language Research; Literacy; Literature; *Measurement Techniques; Oral Language; Reading Instruction; Teacher Attitudes; *Test Reviews; Writing Instruction
IDENTIFIERS	*The Research Instruments Project

ABSTRACT

This volume of measurement instruments in the English language arts is the second one produced by the Research Instruments Project (TRIP), which was designed to collect and evaluate research instruments in language arts, language and language development, literacy, literature, oral language, reading, teacher knowledge/attitudes, and writing. The 160 measurement instruments are arranged alphabetically by category; within each category, measures are listed alphabetically by author. All instruments are cross-referenced by author. The age range indicated is the specific age grouping as stated by the authors or the age of the sample to whom the instrument was administered. The description of the instrument provides the purpose of the instrument, the date of construction, and a physical description of the instrument--often including sample items and administration data (directions, time, scoring procedures, and so forth). The lack of reliability and validity data for instruments is indicated when the information was unavailable. In the case of tables of difficulty, indexes, and so on, the data contained are summarized and the complete data are made available with the test or references cited. (HOD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED255933

MEASURES FOR RESEARCH AND EVALUATION IN THE ENGLISH LANGUAGE ARTS

VOLUME 2

WILLIAM T. FAGAN JULIE M. JENSEN CHARLES R. COOPER

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

CS208842

ERIC®
NCTE 2

This work is dedicated to Bernard O'Donnell, who played a most significant role in the publication of Volume 1 and who nurtured the goal for Volume 2.

Measures for Research and Evaluation in the English Language Arts, Volume 2

William T. Fagan
University of Alberta

Charles R. Cooper
University of California

Julie M. Jensen
University of Texas

ERIC Clearinghouse on Reading and Communication Skills
National Institute of Education

National Council of Teachers of English
1111 Kenyon Road, Urbana, IL 61801

NCTE Editorial Board: Candy Carter, Delores Lipscomb, John S. Mayher, Elisabeth McPherson, John C. Maxwell, *ex officio*, Paul O'Dea, *ex officio*.

Consultant Readers: Kenneth Kantor and Alan Purves.

Book Design: Tom Kovacs for TGK Design

NCTE Stock Number 31018

Published 1985 by the ERIC Clearinghouse on Reading and Communication Skills and the National Council of Teachers of English, 1111 Kenyon Road, Urbana, Illinois 61801.

This publication was prepared with funding from the National Institute of Education, U.S. Department of Education, under contract no. 400-83-0025. Contractors undertaking such projects under government sponsorship are encouraged to express freely their judgment in professional and technical matters. Prior to publication, the manuscript was submitted to the Editorial Board of the National Council of Teachers of English for critical review and determination of professional competence. This publication has met such standards. Points of view or opinions, however, do not necessarily represent the official view or opinions of either the National Council of Teachers of English or the National Institute of Education.

Contents

Preface vii

Introduction xi

Measures

- Language Arts 1
- Language and Language Development 5
- Literacy 23
- Literature 27
- Oral Language 59
- Reading 65
- Teacher Knowledge/Attitudes 133
- Writing 161
- Miscellaneous 223

Appendix 229

Indexes

- Author 235
- Category 240

Preface

In 1975, *Measures for Research and Evaluation in the English Language Arts* was published by the National Council of Teachers of English and the ERIC Clearinghouse on Reading and Communication Skills. Five years later, about 4,000 copies had been sold and the book had gone into its second printing. Since the book was aimed at a very specific group of users, this was considered to be a very good record of its success. Based on information on use and on an awareness of the continued construction of measures for evaluating various aspects of the English language arts, it was decided to make available a selection of these measures in *Measures for Research and Evaluation in the English Language Arts, Volume 2.*

The most current instrument in Volume 1 was constructed in 1973. During the years since then, a number of changes have occurred within the English language arts that are reflected in the instruments being used for research and evaluation. One change is due to a greater emphasis on developing theoretical positions to provide a framework for the interpretation of various data, and the resulting availability of a number of theoretical positions in any one field.

In the field of reading, two particular areas have taken on a new emphasis. Even though early writers (Huey 1908; Thorndike 1917) suggested studying reading as a cognitive activity, this suggestion largely went unheeded and the delineation and classification of skills became the focus for well over half the century (Gray 1917; Davis 1944; Thurstone 1946; Cleland 1965; Barrett 1968). The rise of transformational grammar and the emergence of the psycholinguistic approach to reading redirected the focus to analyzing what readers do when they read. However, the initial concern was usually with the smaller units of linguistic information. Errors in oral reading were often viewed as deviant perceptual forms (Nicholson 1977), though there were implications for comprehension. Syntax was the main concern of transformational/generative grammarians and the sentence was viewed as the most viable unit for analysis. Certain researchers, however, expounded on the broader view of reading as a language activity (Goodman 1968) and on reading as a constructive rather than a reproductive process (Neisser 1967). Within recent years, many instruments have appeared that have focused on describing more precisely the nature of this constructive process. Concurrent with this movement was support for the view that readers operated on units larger than the sentence and that meaning could not be divorced from comprehension, memory, and learning. Earlier works (Bartlett 1932) were being reexamined in a new light to help provide answers to how readers remembered longer passages, and various conceptual frameworks for analyzing text structure (Kintsch and Monk 1972; Mandler and Johnson 1977; Meyer 1975; Rumelhart 1977; Stein 1978; Thorndyke 1977) and knowledge structure (Adams and Collins 1979; Anderson et al. 1977; Fredericksen 1975; Shank 1975) were popularized. Consequently, a whole new set of analytic and descriptive instruments became available.

Changes in research on the production of written language tended to parallel the changes occurring within the field of reading. Emphasis on the evaluative aspects of writing declined and instead the focus centered on describing the processes that writers engage in prior to, during, and following the writing act. Again, as in the case of reading, the total output, as opposed to sentences or other smaller units, was the focus of attention. The emphasis on describing writing can also be seen by its influence on research in teacher knowledge and attitudes. One half of the instruments in this category which were reviewed for inclusion in this publication were concerned with teachers' attitudes toward writing instruction, toward themselves as writers, and toward the processes and skills of writing.

The changes in the research in language and language development, including children's meta-knowledge of various aspects of language, were also dominated by description rather than evaluation. Because of the emphasis on describing children's language behavior in natural contexts, the research does not lend itself to devising "instruments" as defined for this project. Consequently, the number of instruments available for studying the various meta-aspects (meta-knowledge, meta-memory, metacognition, etc.) does not reflect the growing amount of research in this area.

A second change, that no doubt is influenced by the first, deals with the shifting of emphasis among the different areas of the English language arts. In the years since the publication of Volume 1, research in the field of writing has made greater progress than similar work in any of the other fields, with reading research following second, perhaps due largely to the work on discourse processing. The emphasis that has been given to each of these fields is reflected in the number of instruments for each of these areas.

This volume of *Measures for Research and Evaluation in the English Language Arts* would not have been possible without the interest and support of the sponsoring organizations: the NCTE Research Foundation, the National Council of Teachers of English, and the ERIC Clearinghouse on Reading and Communication Skills. Individuals who should be singled out for recognition include the late Bernard O'Donnell, who fostered an interest in the publication of Volume 2, and who coordinated the involvement of the ERIC Clearinghouse. Fran Biederman and Dorcas Rohn arranged for the dissemination of information about the project, forwarded submissions to the committee members, and assisted with the preparation of the final manuscript. Holly O'Donnell was responsible for setting up a computer search of the ERIC system to locate instruments to be considered for inclusion. The committee members reviewed the submissions and searched out others that were considered to have the potential for continued research. The committee is responsible for the information as it is presented.

W. T. Fagan
C. R. Cooper
J. M. Jensen
The Research Instruments
Project Committee

Introduction

The availability of information and its use are two different issues. With the publication of Volume 2, approximately 160 measures for research and evaluation will have been made available for potential users. It seems that the bulk of research is done by one-time researchers—the graduate students (Weintraub and Farr 1976). Consequently, it is expected that the information in Volumes 1 and 2 may be utilized most heavily by this group. The committee recognizes that many researchers did not respond to our invitation to submit instruments and this may have resulted in few instruments in some categories and the omission of important instruments in others.

The text could be used as part of a research seminar for graduate students since their concerns tend to apply to the basic rather than the esoteric aspects of research. Prospective researchers could well benefit by discussing instrument designs in terms of their purpose, their range of use, and their validity and reliability. In Volume 1, ten suggestions were listed for researchers who may be constructing instruments. It is interesting that the implementation of many of these suggestions (especially validity and reliability) was not always evident in the instruments reviewed for this volume. These suggestions, which are reprinted at the end of this section, could be used by graduate students to evaluate the instruments presented here and to serve as a checklist for evaluating instruments that they themselves might develop. It should be noted that the instruments included here do not all follow the "classical empirical design" utilizing statistical procedures. Many of the instruments are suitable for studies of a descriptive or ethnographic nature.

In addition to the use suggested above, the volume may also be consulted by graduate students seeking dissertation topics. Since the vast majority of these instruments grew out of research studies, there is always a need to replicate or expand the initial research. The references listed can provide a broader framework to aid a prospective researcher in conceptualizing a research problem. If the instrument reviewed here is to be used, then the great advantage is that *the complete instrument is available,* either in this volume or in the ERIC system.

Furthermore, it is hoped that the contribution of this volume in assisting with research plans will extend beyond the neophyte, since the continuing researcher is always eager to sharpen his or her research skills and may benefit in the ways suggested previously. A perusal of this volume will give the researcher a sense of the direction that research is going and consequently should suggest gaps that need to be filled.

Format of Volume 2

In consideration of those researchers who have used Volume 1 and to aid those who may need to refer to Volume 1 (having encountered Volume 2 first), it was decided to pattern Volume 2 as closely as possible after Volume 1 so that information from both might be accessed more easily. At one point, the committee considered reproducing some of the instruments from Volume 1 in Volume 2, but this soon became an impossible task since all the instruments in Volume 1 may be considered equally valuable depending on a researcher's particular needs. Procedures for collecting and evaluating data and the guiding definitions were the same as those used in Volume 1 and will not be reprinted here. The category index to Volume 1 is included in the Appendix, which begins on page 229.

Organization of the Report

For those who have not used Volume 1, specific information needed to interpret the report is given below, with any changes from the first volume noted.

Arrangement. Measurement instruments are arranged alphabetically by category; within each category, measures are listed alphabetically by author. The categories are language arts, literacy, language and language development (including listening, speaking, and meta-studies), literature, oral language, reading, teacher knowledge/attitudes, and writing. The purpose of the instrument determines its assignment to a category. All instruments are cross-referenced by author. When there is more than one author, the name of the first author, considered the primary author, is stated.

Age Range. It was decided not to use only the categories established for Volume 1 (Preschool, Primary, Intermediate, Junior High, Senior High, Postsecondary–adult). Instead, the specific age grouping as stated by the authors was also used. If the instrument was obtained from a research report, then the age of the sample to whom the instrument was administered was used in deciding the age-range category.

Description of Instrument. Under this heading the following data are included: the purpose of the instrument, the date of construction, and a physical description of the instrument, often including sample items and administration data (directions, time, scoring procedure, etc.).

In some instances, desired data were not available. For example, it might be assumed that the date of construction of an instrument used in a doctoral study was the same as the completion date of the dissertation. When the date of construction was not clear, the completion date of the dissertation or publication date of a journal article was used to give some guide to the potential users as to the recency of the instrument. Time for administration was not always given and it would have been hazardous for the committee to assign an approximate time. When such data are not available, it is hoped that the description, including the number of items and the directions, might help the potential user assess whether this instrument is useful for his or her purposes. Brief directions were included in full, while lengthy directions were summarized. If the test was short, it was included here in full; if not, at least one sample item was given.

Validity, Reliability, and Normative Data. Questions of reliability and validity are crucial ones for most of the measures included in this monograph. The reader might wish to review current knowledge on these two topics, and the authors recommend Anastasi's *Psychological Testing.* Her chapter on reliability is brief but adequate, and her two chapters on validity constitute a comprehensive and readable introduction to that topic. The lack of reliability and validity data for instruments was indicated when the information was unavailable. In the case of tables of difficulty indexes and so on, the data contained therein were summarized and the complete data were made available with the test or the reference(s) cited.

Ordering Information. The main source is the ERIC Document Reproduction Service, Computer Microfilm International Corporation, 3900 Wheeler Avenue, Alexandria, Virginia 22304. References bearing ERIC Journal (EJ) numbers appear only in an annotated format; those having ERIC Document (ED) numbers are available in their entirety. A second source for some items is University Microfilms, Dissertation Copies, 300 North Zeeb Road, Ann Arbor, Michigan 48106. Pencil and paper portions of the test may be obtained on microfiche or microfilm from the above sources. If the instrument includes hardware, the source of this equipment (as supplied by the author) is indicated.

Related Documents. The key reference to the instrument is listed. In the case of doctoral dissertations, this is the dissertation itself. A research instrument has been modified from several sources in some instances; in such cases, all sources are listed.

Suggestions for the Construction of Research Instruments

1. State specifically the purpose for which the instrument is being constructed. A stated purpose such as "to measure reading comprehension" will not reveal to the reader that the test was actually designed to measure inferential reasoning in the expository material of fourth-grade children.
2. Indicate clearly the steps followed in constructing the instrument, including such things as the directions given to judges who evaluated test items.
3. Establish validity for the test.
4. Establish reliability for the test.
5. Give the test a title and be consistent in using this title when referring to the test.
6. State the theoretical framework that gives meaning to the test and to which the test results add further information.
7. State the assumptions underlying what the test proposes to measure.
8. Describe the sample to whom the test was administered and on whom the validity and reliability data may be obtained.
9. If at all possible, include the complete tests in the research report or doctoral dissertation, or state where the instrument may be obtained.
10. Share the measure with other potential users. Cooperation in responding to brochures such as the one that solicited instruments for this project is appreciated.

W. T. F.
C. R. C.
J. M. J.

References

Adams, M. J., and A. Collins. "A Schema-Theoretic View of Reading." In *New Directions in Discourse Processing,* edited by R. O. Freedle. Norwood, N.J.: Ablex Publishing, 1979.

Anastasi, A. *Psychological Testing.* 3d ed. New York: Macmillan, 1969.

Anderson, R. C., R. J. Spiro, and M. C. Anderson. *Schemata as Scaffolding for the Representation of Information in Connected Discourse.* Technical Report no. 24. Urbana: University of Illinois, Center for the Study of Reading, 1977.

Barrett, T. C. "Taxonomy of Cognitive and Affective Dimensions of Reading." In *What is reading?* edited by T. Clymer. Innovation and Change in Reading Instruction. 67th Yearbook of N.S.S.E. Chicago: University of Chicago Press, 1968, 1–30.

Bartlett, F. C. *Remembering.* Cambridge, Eng.: Cambridge University Press, 1932.

Cleland, D. L. "A Construct of Comprehension." In *Reading and Inquiry,* edited by J. A. Figurel. Newark, Del.: International Reading Association, 1965.

Davis, F. B. "Fundamental Factors of Comprehension in Reading." *Psychometrika* 9 (1944): 195–97.

Fredericksen, Carl. "Representing Logical and Semantic Structure of Knowledge Acquired from Discourse." *Cognitive Psychology* 7 (1975): 317–458.

Goodman, Kenneth. *The Psycholinguistic Nature of the Reading Process.* Detroit: Wayne State University Press, 1968.

Gray, Clarence T. *Types of Reading Ability as Exhibited through Tests and Laboratory Experiments.* Supplementary Educational Monographs, no. 5. Chicago: University of Chicago Press, 1917.

Huey, E. B. *The Psychology and Pedagogy of Reading.* New York: Macmillan, 1908.

Kintsch, W., and D. Monk. "Storage of Complex Information in Memory: Some Implications of the Speed with Which Inferences Can Be Made." *Journal of Experimental Psychology* 94 (1972): 25–32.

Mandler, J. M., and N. S. Johnson. "Remembrance of Things Parsed: Story Structure and Recall." *Cognitive Psychology* 9 (1977): 111–51.

Meyer, B. J. *The Organization of Prose and Its Effect upon Memory.* Amsterdam: North Holland Publishing, 1975.

Neisser, U. *Cognitive Psychology.* New York: Appleton-Century-Crofts, 1967.

Nicholson, Tom. "The Relative Effects of Different Error Types on Understanding Connected Discourse." Ph.D. diss., University of Minnesota, 1977.

Rumelhart, D. E. "Understanding and Summarizing Brief Stories." In *Basic Processes in Reading: Perception and Comprehension,* edited by D. Laberge and S. J. Samuels. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1977.

Shank, R. C. "The Structure of Episodes and Memory." In *Basic Processes in Reading: Perception and Comprehension,* edited by D. Laberge and S. J. Samuels. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1975.

Stein, N. L. *How Children Understand Stories: A Developmental Analysis.* Urbana: University of Illinois, Center for the Study of Reading, 1978.

Thorndike, Edward L. "Reading as Reasoning." *Journal of Educational Research* (1917): 323–32.

Thorndyke, P. W. "Cognitive Structures in Comprehension and Memory of Narrative Discourse." *Cognitive Psychology* 9 (1977): 77–110.

Thurstone, L. L. "Notes on a Reanalysis of Davis' Reading Tests." *Psychometrika* 11 (1946): 185–88.

Weintraub, S., and R. Farr. "Introduction." In *Improving Reading Research,* edited by R. Farr, S. Weintraub, and B. Tone. Newark, Del.: International Reading Association, 1976.

Errata for Volume 1

The committee regrets the omission of certain data from Volume 1. The following references were inadvertently omitted from the "Introduction":

Gunderson, D. V. "Flaws in Research Design." *Research in the Teaching of English* (Spring 1967): 10-16.

Petty, W., P. Herold, and E. Stoll. *The State of Knowledge about Teaching Vocabulary.* Champaign, Ill.: National Council of Teachers of English, 1968.

Purves, A. C., and R. Beach. *Literature and the Reader: Research in Response to Literature, Reading Interests and the Teaching of Literature.* Urbana, Ill.: National Council of Teachers of English, 1972.

Sherwin, J. S. *Four Problems in Teaching English: A Critique of Research.* Scranton, Penn.: International Textbook Co., 1969.

Summers, E. "Reading in the Secondary School." *Review of Educational Research* 37 (April 1967): 134-51.

Our apologies to Dr. Carol Sager, author of the Sager Writing Scale (Volume 1, 203) for the misinformation on reliability data for children using the scale. The correct information should have been: To determine whether children should be taught to use the Revised Scale with some degree of proficiency, an estimate of reliability for the Revised Scale and for each of its four components was computed by means of an interclass correlation. Ten stories and the ratings of ten children, who were randomly selected, were used. The estimated reliability of the four components of the Revised Scale ranged from .96 to .98, with a total estimated reliability of .99.

The wording of all extracts appears as originally submitted.
ERIC/RCS

Language Arts

Category: Language Arts

Title: Curriculum Review Handbook: Language Arts

Author: Oklahoma State Department of Education

Age Range: Kindergarten through twelfth grade

Description of Instrument:

Purpose: Based on the rationale that a language arts program should have sequential steps for developing language skills that spiral from one level to another, this handbook is designed to provide general goals for evaluating language arts programs at any level.

Date of Construction: 1981

Physical Description: The major section presents a checklist that evaluates the current program according to its general and specific goals. The objectives used are based on those recognized by the National Council of Teachers of English to reflect basic skills. While the general checklist is arranged according to the headings—philosophy, curriculum, facilities/materials, teacher involvement, and student involvement—the specific checklist focuses on skills in six divisions from kindergarten to secondary/high school. A comparison chart is provided indicating the highest possible score for each section of the checklist and against which a particular program may be evaluated.

In addition to the major section, additional sections help educators develop a philosophy and rationale for the language arts program, examine the student outcomes expected, compile listings of the skills and concepts taught in all courses, list resources available to all teachers, list the needs and recommendations that have become apparent during the evaluation, develop a checklist of essential skills for all levels of the program, and develop course descriptions.

Validity, Reliability, and Normative Data: Content validity may be claimed for the checklist since the objectives used are based on those recognized by the National Council of Teachers of English to reflect basic skills.

Ordering Information: ED 208 540

Language and Language Development

Category: Language and Language Development

Title: Interactional Competency Checklist (ICC)

Author: Janet K. Black

Age Range: Preschool to adult

Description of Instrument:

Purpose: To assess ability to interact in communicative contexts.

Date of Construction: 1978

Physical Description: This instrument is a tool for observing, documenting, and analyzing oral language samples. It was initially designed to assess kindergarten children's interactional competence; however, it can be adapted for use with any age and in any conversational context. The checklist follows, excluding spaces under each category to record examples and dates.

Interactional Competency Checklist

I. Ability to Adapt to Changes in the Setting

This category is based on Cicourel's second and fourth properties of interactional competence and attempts to assess whether young children can "behave as if they share the same social setting and are receiving and processing the same information," and "can normalize discrepancies to sustain social interaction" (Cicourel 1972, 217–18). These properties were adapted to the sociodramatic environment under the following subcategories:

A. Adjusts to the various themes of play (or conversation) ☐

B. Extends the organization of the plot (or conversation) ☐

C. Extends character development ☐

II. Nonverbal Appropriateness

This category is based on Cicourel's fourth property of interactional competence, and attempts to assess whether young children possess "'normal form repertoires' of possible appearances, behaviors, and utterances that can be understood when emergent in contextually organized settings" (Cicourel 1972, 217–18). This property was adapted to the sociodramatic environment under the subcategories of gestures, facial expressions, and body movements. In addition, two other subcategories—nonverbal behavior and vocal intonation and stress—were included based upon Mishler's (1976) research that indicated first graders' use of appropriate stress and intonation.

A. Uses appropriate gestures ☐
B. Uses appropriate facial expression ☐
C. Uses appropriate body movement ☐
D. Uses appropriate vocal intonation ☐
E. Uses appropriate stress ☐

III. Familiarity with Normal Constraints and Conditions

This category is based upon Cicourel's third property of interactional competence, which concerns such items as a knowledge of who can speak first or next, what topics are considered socially relevant, how to terminate an exchange, repair, recycling, and repeating (Cicourel 1972, 217–18). These subcategories have been investigated in adults (Schefhoff and Sacks 1973, 1974) and in children (Mishler 1976).

A. Knows when to speak first or next ☐
B. Discusses topics socially relevant to the situation ☐
C. Knows how to terminate a conversation ☐
D. Repairs (corrects) oral language ☐
E. Recycles (rephrases) oral language ☐
F. Repeats oral language ☐
G. Uses interrogatives ☐

IV. Sequencing

This category is based upon Cicourel's first and sixth properties of interactional competence and attempts to assess whether young children demonstrate the ability to think back or reflect upon previous future informational events, objects, and resources within the communicative setting (Cicourel 1972, 217–18).

A. Links past experience with present information events ☐
B. Links past experience with possible future informational events ☐

Validity, Reliability, and Normative Data: Communicative competence consists of grammatical competence and interactional competence. Since it is ultimately the children's ability to use appropriate language in particular communicative settings that determines their communicative competence, oral language assessment must include observing children's interaction skills in various social contexts. The ICC provides guidelines for assessing children's interactional abilities.

This instrument was based on Hymes's rationale, which suggests that as children attend to social interaction and participate in it with family members and significant others they learn communicative competence. Hymes contends that normal children acquire knowledge of sentence structure not only as it is grammatical but also as it is appropriate. Cicourel seems to refer to the same process but calls it "interactional competence." He suggests that there are seven properties to be utilized if one is to be interactionally competent. These properties have been adapted to the various ICC categories.

The ICC as a method of evaluation considers: (1) the issues of cultural, performative, and teacher strategy biases that exist in standardized test instruments; (2) the problem of artificiality inherent in the interaction of test administrator and subject; and (3) the research pertaining to the nature of language. The author observed the twelve students of a kindergarten class in a private school as they engaged in spontaneous dramatic activities. Examples of the subjects' interactional competence were documented using the ICC checklist. The study concluded that observing children's language in an informal context generally provided more comprehensive information than did two formal oral language assessment instruments used in the study.

Ordering Information: Full instrument reproduced above.

Related Documents:

Black, Janet K. "Formal and Informal Means of Assessing the Communicative Competence of Kindergarten Children." *Research in the Teaching of English* 13 (February 1979): 49-68. (EJ 204 423)

———. "There's More to Language Than Meets the Ear: Implications for Evaluation." *Language Arts* 56 (May 1979): 526-33. (EJ 204 612)

Category: Language and Language Development

Title: Diagnosis of Language Competency Inventory (DLCI)

Authors: Howard E. Blake
Ethel M. Maull

Age Range: Primary

Description of Instrument:

Purpose: To measure children's receptive and expressive language competence.

Date of Construction: 1977

Physical Description: Designed to be administered individually, this fifty-seven-item instrument inventories children's performances in the following eight language subareas: motor functions, memory functions, visual functions, tactile-kinesthetic functions, vocal functions, auditory functions, following directions, and language concepts. Each item has been given an arbitrary weight. Eight subscores and a total score are recorded on either an individual scoring sheet, a group scoring sheet, or both, depending on the purpose of the testing.

The partial individual recording sheet shown on page 11 lists the nine items in the "Motor Functions" category. In item 1.1, for example, the child is asked to hop and to bounce a ball. In 1.3 the child is to mime hair combing, teeth brushing, and hand washing.

In subarea two, "Memory Functions," the child is presented with four objects or pictures. One is then hidden and the child must identify the missing one. Subarea three, "Visual Functions," includes association tasks (match, name, or find an object or shape) and closure tasks (complete a picture, letter, numeral, word, or sentence). "Tactile-Kinesthetic Functions," category four, include the ability to identify common objects with eyes closed. "Vocal Functions" are scored for quantity of output, completeness, ease of expression, vocabulary, and grammar as the child talks about the subjects "What do you like to do after school?" and "What was the funniest thing that ever happened to you?" "Auditory Functions" include identifying sounds, discriminating sounds, and recalling patterns of sounds. In "Following Instructions," the child must carry out single, double, and triple commands. Finally, "Language Concepts" include recognizing common objects, knowing uses of objects, recognizing colors, understanding *yes* and *no,* knowing a spatial vocabulary, and responding appropriately to quantity, size, and value terms.

Validity, Reliability, and Normative Data: No reliability or validity data are reported for this instrument, though a four-year process of develop-

Individual Recording Sheet

Name ______________________

		Scores					
Checkpoint Dates							
1. Motor Functions (11 points)							
1.1 Gross motor tasks	(1)						
1.2 Imitating	(1)						
1.3 Pantomiming	(1)						
1.4 Design copying	(1)						
1.5 Figure drawing	(2)						
1.6 Letter copying	(1)						
1.7 Writing name	(1)						
1.8 Numeral copying	(1)						
1.9 Sentence copying	(2)						

ment and refinement is noted, during which the inventory was used with children in beginning reading programs, in regular and special education classrooms (particularly with brain-damaged children). The authors point to current research on oral/print relationships and on self-image. Their objective is to support teachers working to identify strengths, remedy weaknesses, and develop appropriate programs. To this end, they urge that items and weighting be adapted to the children and the program involved, thereby improving content validity.

Ordering Information: ED 236 650

Related Document:

Blake, Howard E., and Ethel M. Maull. "Diagnosing Language Competency." Paper, Temple University, 1977.

Category: Language and Language Development

Title: Comprehension of Ambiguous and Other Polysemous Utterances Presented in Spoken Mode

Author: Rita S. Brause

Age Range: Age seven through adult

Description of Instrument:

Purpose: To assess listeners' comprehension of ambiguous and polysemous utterances.

Date of Construction: 1975

Physical Description: This instrument is a series of twenty-four statements that includes varying degrees and types of ambiguity. Each of three sets of eight items includes, in the order listed, the following types of ambiguity: (1) one meaning of a word in an unambiguous context, (2) a second meaning of the same word in an unambiguous context, (3) multiplicity of pronoun referents, (4) figurative language, (5) symbolic language of proverbs, (6) lexical ambiguity, (7) surface structure ambiguity, and (8) deep structure ambiguity. In individual interviews, subjects listen to "endless" prerecorded audiotapes of these statements, presented in a neutral tone that provides no stress or pitch cues to aid disambiguation. Subjects are asked to explain each statement. Only those interpretations similar to paraphrases agreed upon by validation experts are considered correct. Following is the first set of eight statements:

1. The boy is a member of the chess club.

   Paraphrases:
   X Club means a group of people.
   Club means a stick or a weapon used for protection.

2. The policeman brought a heavy club for protection when going to stop the fight.

   Paraphrases:
   Club means a group of people.
   X Club means a stick or a weapon used for protection.

3. Mickey knew that he was going to be late (C. Chomsky 1969).

   Paraphrases:
   X Mickey Mouse knew that Mickey himself was going to be late.
   X Mickey Mouse knew that someone else (Donald Duck) was going to be late.
   Mickey is always late.

4. Aunt Helen spilled the beans.

   Paraphrases:
   - Aunt Helen likes to cook beans.
   - X Aunt Helen told someone's secret.
   - X Aunt Helen allowed the beans that she was preparing to fall out of the pot or the dish.

5. Little streams make mighty rivers (Piaget 1955).

   Paraphrases:
   - X Small streams of water may become mighty by becoming part of rivers.
   - Small amounts of water help the river to grow.
   - X Small or young people may grow to be important or influential adults or grown-ups.

6. The policeman took his club to the riot.

   Paraphrases:
   - X The policeman carried his own stick to a riot.
   - The policeman is attending a meeting of his group.
   - X The policeman brought his fellow club members to a fight or riot.

7. They fed her dog biscuits.

   Paraphrases:
   - X A girl is being given biscuits to eat which are usually eaten by dogs.
   - X A girl's dog is being given biscuits to eat.
   - A dog is being given some meat to eat.

8. The mayor will ask the police to stop drinking.

   Paraphrases:
   - X The mayor will ask the police to stop their own drinking.
   - The mayor visited all the bars in the city.
   - X The mayor asked the police to stop other people from drinking.

X denotes paraphrases considered acceptable.

Validity, Reliability, and Normative Data: Content validity for statements used is based on their selection from theoretical and empirical sources such as Piaget and C. Chomsky. Interrater reliability for scoring answers was established at the .96 level. A Guttman scaling of reproducibility of findings in the hierarchy of comprehension yielded a coefficient of .97. A hypothesized hierarchy of acquisition was developed based on the findings and related in part to age and educational background:

Stage One: Understanding of two meanings of a word when presented in unambiguous contexts.

Stage Two: Understanding lexical ambiguity.
Understanding figurative language.
Stage Three: Understanding surface structure ambiguity.
Understanding multiplicity of pronoun referents.
Stage Four: Understanding deep structure ambiguity.

Ordering Information: ED 236 666

Related Documents:

Brause, Rita S. "Developmental Aspects of the Ability to Understand Semantic Ambiguity, with Implications for Teachers." *Research in the Teaching of English* 11 (Spring 1977): 39–48. (EJ 160 312)

———. "On Ambiguity: An Investigation of the Development of the Ability to Understand Semantic Ambiguity." Ph.D. diss., New York University, 1975. *Dissertation Abstracts International* 36 (1975): 2173-A.

Category: Language and Language Development

Title: Comprehension of Ambiguous and Other Polysemous Utterances: Presented in Written Mode

Author: Rita S. Brause

Age Range: Ages ten through sixteen

Description of Instrument:

Purpose: To assess comprehension of written ambiguous and polysemous structures.

Date of Construction: 1979

Physical Description: Three items assess comprehension of each of the following structures: (1) one meaning of a word in an unambiguous context, (2) a second meaning of the same word in an unambiguous context, (3) multiplicity of pronoun referents, (4) figurative language, (5) symbolic language of proverbs, (6) lexical ambiguity, (7) surface structure ambiguity, and (8) deep structure ambiguity. This written test assesses comparable structures and follows the same design of presentation as "Comprehension of Ambiguous and Other Polysemous Utterances: Presented in Spoken Mode" found elsewhere in this volume. In the first part of the test students are asked to identify semantically acceptable paraphrases for an initial sentence:

Mom, we finished all of the soda.
_____ a. We drank up all the soda.
_____ b. There isn t any soda in the house to drink.
_____ c. Where is the soda?

Then students must independently paraphrase in writing all of the meanings in the same set of sentences. In the final part of the test, students are asked to independently paraphrase nine sentences, e.g., "Most people wear clean socks and shoes." Comprehension of a structure is demonstrated by two acceptable interpretations out of the three items testing each aspect of ambiguity.

Validity, Reliability, and Normative Data: The hypothesized sequence of development (Brause 1975, 1977) of a listener's ability to comprehend ambiguous and polysemous utterances based on the findings assessing comprehension in a spoken mode was supported in this study; there were no invariant cases. Content validity is based upon the selection of sentences from theoretical and empirical sources such as Piaget and C. Chomsky. An interrater reliability coefficient of .94 was obtained by comparing the scores of two independent judges on five participants' responses.

Ordering Information: ED 236 663

Related Documents:

Brause, Rita S. "Comprehension of Ambiguous Sentences Presented in Spoken and Written Modes." Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, California, April 1979.

———. "Developmental Aspects of the Ability to Understand Semantic Ambiguity, with Implications for Teachers." *Research in the Teaching of English* 11 (Spring 1977): 39–48. (EJ 160 312)

———. "On Ambiguity: An Investigation of the Development of the Ability to Understand Semantic Ambiguity." Ph.D. diss., New York University, 1975. *Dissertation Abstracts International* 36 (1975): 2173-A.

Category: Language and Language Development

Title: Listening Skills Assessment

Author: Thomas G. Devine

Age Range: Grade 6 (abbreviated form), grade 11 (full test)

Description of Instrument:

Purpose: To assess listening ability and to indicate implications for listening instruction.

Date of Construction: 1980

Physical Description: The full form of this test contains forty-five items that measure fifty-three specific listening skills from the following five categories: simple recall, recognizing and following spoken directions, recognizing a speaker's purpose and plan, critical listening, and higher-level listening skills. The abbreviated test form consists of the first thirty items. The test requires no reading. Its multiple-choice items are developed around these "real life" listening situations: conversations overheard on the street or in the supermarket, talk at meetings, radio weather forecasts, spoken directions given by a service station manager, and semi-formal talks which might be heard on radio or television. The full version of the test takes 45 minutes to complete. Students respond by filling in slot A, B, C, or D on a machine-scoreable answer sheet. Following is one listening situation and one of the eight items developed from it:

> In this third part of this listening assessment, you, the listener, are supposed to be working part-time in a local service station. The manager is extremely busy and gives you these directions in a hurry.
>
> Manager: Go to the back room. (The key is in the top left-hand drawer of the office desk.) Get three 5-gallon cans of that new synthetic motor oil. Put them on the edge of the back drive because Henry will pick them up. Then run down the street to the Mobil station and tell Ted that I need the wrench he borrowed from me yesterday. (Ted is the big guy with the mustache.) Give the wrench to Sam in the shop as soon as you get it and then relieve me at the pump. But before you go, don't forget to put the oil out.
>
> Now, let's see how well you listened to these directions.
>
> Question 13. Where is the key?
> A. in the top left-hand drawer
> B. on a hook in the hall
> C. in the service truck
> D. in the manager's pocket

Validity, Reliability, and Normative Data: Content validity was established by searching the research literature to compile a master list of listening skills. From the master list of more than forty skills, eight were selected for inclusion. After field-testing, editing, and refining items, the eight skills were placed into the five final, more general categories. Several dozen listening situations were created and shared with students from two high schools. From comments and suggestions made by the students it became clear that certain situations worked while others did not. A preliminary test was prepared based on these discussions and was piloted on 120 grade 11 students from two high schools. The instrument and its items were also studied by a panel of authorities in the field. An item analysis was performed, revealing that thirty-six items were acceptable. Thirty-four of these were retained in the final version and eleven new items were generated. Reliability was determined by the split-half method, which yielded a coefficient of .76.

Ordering Information: ED 236 651

Related Document:

Devine, T. G. "Listening: What Do We Know after Fifty Years of Research and Theorizing?" *Journal of Reading* 21 (1978): 296–304. (EJ 169 540)

Category: Language and Language Development

Title: Techniques for Collecting Literacy Events from Young Children

Author: Yetta M. Goodman

Age Range: Ages two through seven, and with readers who seem to be having difficulty.

Description of Instrument:

Purpose: To reveal children's awareness of print and use of contextual clues. To discover metalinguistic awareness, language use, and attitudes toward reading and writing.

Date of Construction: 1980

Physical Description: The instrument is comprised of six tasks. The first three are print awareness tasks designed to reveal children's awareness of print and use of supporting context clues. Familiar food, toy, and household product labels, or store and street signs are mounted on tagboard and presented to children. The tasks are gradually contextualized. In the first task, the stylized logo is used in full color, surrounded by the familiar symbols, pictures, and designs associated with the items. The same materials are used in task two. The print retains the stylization and color but is not accompanied by surrounding designs, pictures, or symbols. In task three, the same logos, trademarks, and signs used in tasks one and two are written in black type on white index cards. This level of decontextualization lacks any supporting color, picture, symbol, or design cues. The fourth and fifth tasks are aimed at discovering the children's metalinguistic awareness, language use, and attitudes toward reading and writing. These tasks are in an interview format. A writing sample is also collected during task five. The sixth task, a book handling knowledge task, is designed to reveal children's knowledge about and use of print in books.

> Tasks 1, 2, and 3:
>
> Sample Items: Campbell's Tomato Soup, Lego, Crest Toothpaste, Chicken of the Sea Tuna, McDonald's, Johnson and Johnson Baby Powder, School Bus, Stop, Coca-Cola, Chevrolet, Ivory Soap, Phone, School, Kellogg's Rice Krispies, Kentucky Fried Chicken, and Gerber's Baby Food.
>
> Sample Questions: Have you ever seen this before? Where? What do you think that is? What do you think that says? How do you know? Why did you say that? What tells you that it says ______ ? Show me with your fingers where it says ______ .
>
> Task 4: (fourteen interview questions):
>
> Sample Question: Do you know how to read? (if "yes," How did you learn to read? Did somebody help you to learn? [if

"yes," who?] or did you learn by yourself? Do you like to read? What do you like to read?)
(if "no," Do you want to be able to read? How will you learn to read? Does someone have to help you learn how to read? Who do you think will help you learn how to read?)

Task 5: (twelve writing-related tasks)

Sample Task: Write for me. (If no response, say "Pretend to write for me.") Why did you choose those? (paper and writing implements)

Task 6: (twenty-two items related to book handling)

Sample Item: Administrator presents book to child upside down and backwards and says "Show me the front of this book. Take the book and open it so that we can read it together."

Validity, Reliability, and Normative Data: This instrument was developed over a period of seven years by the author, with the assistance of graduate students Bess Altwerger, Lois Bird, Carol Ewoldt, and Myna Haussler. Influenced by the work of Marie Clay and David Doake, it was further expanded through discussions with researchers Jerome Harste, Carolyn Burke, and Virginia Woodward. Techniques for Collecting Literacy Events from Young Children is a descriptive instrument. It can offer valuable insight into developing literacy provided it is coupled with teachers' and researchers' understanding of the development of written language.

Ordering Information: ED 236 647

Related Documents:

Goodman, Kenneth S., and Yetta M. Goodman. "Learning to Read Is Natural." In *Theory and Practice of Early Reading* Vol. I, edited by L. B. Resnick and C. Weaver. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1979.

Goodman, Yetta M. "A Study of the Development of Literacy in Preschool Children." In *Occasional Papers, Program in Language and Literacy,* edited by L. Bird. Tucson: University of Arizona, 1981.

———. "The Roots of Literacy." In *The Claremont Reading Conference Forty-Fourth Yearbook,* edited by M. P. Douglass. Claremont, Calif.: The Claremont Reading Conference, 1980.

Goodman, Yetta M., and B. Altwerger. "Reading: How Does It Begin?" In *Discovering Language with Children,* edited by Gay Su Pinnell. Urbana, Ill.: National Council of Teachers of English, 1980. (ED 193 645)

Category: Language and Language Development

Titles:
1. The Standard English as a Second Language (S-ESL) Spoken English Test
2. The S-ESL Grammar Test
3. The S-ESL Vocabulary Test

Author: Elray L. Pedersen

Age Range: Primary grades through adult

Description of Instrument:

Purpose: To assess the 1) oral comprehension, 2) grammatical fluency, and 3) vocabulary development of students for whom English is a second language.

Date of Construction: 1978

Physical Description: Two forms of the Spoken English Test are available. Form T is a cassette tape recording of ninety items; Form W is a written test of ninety items to be read aloud by the examiner. Three alternate forms—A, B, and C—are available for both tests. Each form employs a multiple-choice format requiring the selection of semantically and grammatically correct responses. The test does not assume knowledge of grammatical terminology. Each test is of 20 minutes duration and may be machine or hand scored. Example:

"Did you sleep well?"
a) I sleeping well.
b) I slept well.
c) I am sleeping well.

Two levels of the Grammar Test have been developed: Form O tests the grammatical fluency of primary graders; Forms A, B, and C are similar tests for secondary or college and university students. The thirty-five questions on each form require no formal knowledge of grammatical terminology. The test requires 15 minutes to complete and can be machine or hand scored. Example:

"_____ your name for me."
a) Write
b) Wrote
c) Writes
d) Written

Two forms of the Vocabulary Test are available; their names and audiences are comparable to those of the Grammar Test. Each form consists of thirty-seven multiple-choice questions of three types: the student fills in a blank with a syntactically and semantically appropriate choice; the student responds to a question about the meaning, value, or usage of underlined words; or the student matches a brief description with one of the choices provided. Administration of each form requires 15 minutes, and the tests can be machine or hand scored. Items come from survival and daily life contexts. Following is one example of each of the three item-types:

"The ____ told the members of her class goodbye."
a) music
b) statement
c) yellow
d) teacher

"1988 will be a ____ to remember."
a) day
b) month
c) year
d) summer

"One cent is how much money?"
a) a penny
b) a nickel
c) a dime
d) a dollar

Validity, Reliability, and Normative Data: The tests are used in the English as a Second Language Program at Weber State College in Ogden, Utah. Test items are patterned after those appearing on such tests as the Michigan Language Test, the Test of English as a Foreign Language (TOEFL), and the ETS English Language Proficiency Test. No reliability data are available.

Ordering Information: ED 236 648

Related Document:

Pedersen. Elray L. "New Instruments for Assessing Language Proficiency." In *Intermountain Teachers of English to Speakers of Other Languages (ITESOL) Papers* 2 (1981): 18-25.

Literacy

Category: Literacy

Title: Literacy Assessment Battery (LAB)

Author: Thomas G. Sticht et al.

Age Range: Adult

Description of Instrument:

Purpose: To be used as a supplement to the Armed Services Vocational Aptitude Battery (ASVAB) for use as a selection and classification instrument for the military services.

More specifically, the LAB was designed to assess the degree to which the "gap" between auding and reading has been closed. The underlying rationale was that many lower aptitude people are unskilled readers and the ASVAB tests, which demand reading skills for successful performance, cannot distinguish who among the poor readers have poor oral language comprehension skills and who have relatively well-developed oral language skills but have problems with written language. Secondly, many job skills depend more on listening and watching than on reading.

Date of Construction: 1982

Physical Description: The LAB consists of three sections: paragraph, vocabulary, and decoding. Within each section, items are presented orally or in reading, or in both modes simultaneously.

The first section consists of four paragraphs, two of which are read silently. The passages are 150 and 190 words long and are at a ninth-grade level as measured by the FORCAST formula for assessing readability. Subjects are required to answer twelve questions per passage involving recall of facts.

The vocabulary section consists of twenty-eight words, fourteen of which are presented in an auding/reading situation, and fourteen of which are presented in a reading situation. The subjects are expected to choose the correct synonym for each word from among the alternatives. The words are taken from the paragraphs in the first section.

The decoding section also contains paragraphs at a ninth-grade level which are adult-oriented. Subjects must detect and circle mismatches between the words on the page and the words read aloud by the examiner. The passages are simultaneously read and auded and are presented at varying rates from 100 to 250 words per minute.

Validity, Reliability, and Normative Data: Concurrent validity was established for the LAB by calculating Pearson product-moment correlations between the LAB total scores and subscores and the total and subscores on the Nelson-Denny Reading Test, Forms C and D; the Gates-MacGinitie Reading Test, Form D; and the Basic Skills Assessment: Reading. All correlations were statistically significant.

Predictive validity was determined by analyzing the records of 551 people with no prior service who were still in the service two years after entry. The LAB proved to be a good predictor of qualification status, probability of being separated from the service for failure to meet minimal behavioral or performance criteria, probability of such separation occurring within the first six months of service or subsequently, months of service before separation, and highest pay grade received.

Three methods were employed to establish reliability for the test. (1) As a measure of internal consistency, all LAB subscores were correlated with each other and with the total score. All correlations were substantial. For example, of people who performed in the fourth quartile on the reading vocabulary subtest, 82 percent also performed in the fourth quartile on the auding/reading vocabulary subtest. (2) Correlations for each of the test subsections as computed by the Kuder-Richardson 21 formula ranged from .83 to .91. (3) As a third measure of reliability, scores on each of the four vocabulary and paragraph subtests were decomposed into parts, each part being associated with one of the two passages making up that subtest. Pearson product-moment correlations between each set of subscores ranged from .71 to .80.

The LAB was normed on 4,599 service applicants who were tested at twenty-five geographically dispersed Armed Forces Examining and Entrance Stations. LAB score conversion tables were prepared by means of an equipercentile norming procedure. These tables allow the conversion of raw scores on any LAB component or on a LAB total score to percentile equivalents, reading grade levels on the Gates-MacGinitie and Nelson-Denny tests, standard scores on the Basic Skills Assessment and percentiles on the Armed Forces Qualification Test (AFQT), and the General Technical (GT) composite of the ASVAB. As an alternative to the equipercentile method for norming the LAB, regression equations were computed for relating LAB scores to Nelson-Denny and Gates-MacGinitie reading grade levels and to AFQT percentiles.

Ordering Information: ED 236 659

Related Documents:

Sticht, T., and L. Beck. *Experimental Literacy Assessment Battery (LAB).* AFHRL-TR-76-51. Lackland Air Force Base, Tex.: Air Force Human Resources Laboratory, August 1976.

Sticht, T., L. Beck, R. Hauke, G. Kleinman, and J. James. *Auding and Reading: A Developmental Model.* Alexandria, Va.: Human Resources Research Organization, 1974. (ED 097 641)

Literature

Category: Literature

Title: Responding to Literature. Communication Skills. Pennsylvania Comprehensive Reading/Communication Arts Plan (PCRP) Assessment Survey I

Author: Stephen M. Koziol, Jr. for the Pennsylvania Department of Education

Age Range: Adult

Description of Instrument:

Purpose: To serve as a vehicle for individual elementary and secondary school teachers to examine and reflect upon their own curriculum design and instructional practices in the teaching of literature and to serve as a tool for schools involved in evaluating their instructional programs and planning long-range improvements.

Date of Construction: 1982

Physical Description: Responding to Literature is one of five survey packets developed to gather accurate, detailed, comprehensive, and descriptive data about instruction in communication skills in a single classroom or course. Packets on the subjects of literature, written composing, oral composing, language proficiency development, and sustained silent reading were designed for teachers engaged directly in communication skills instruction as well as for their students, the parents of their students, and their administrators. Each of the five surveys has four separate forms (one for teachers, one for students, one for administrators, and one for parents) and each addresses two layers of specificity: responses to aspects of the overall communication skill program and responses to a teacher's practices and policies within a single kind of class situation.

The teacher forms are divided into sections related to distinct components of instruction. The Responding to Literature form has ten main sections: (1) "Background Information," (2) "Types of Inclusion," (3) "Purposes for Literature Study," (4) "Selecting Literature for Study," (5) "Pre-reading and Pre-viewing Practices," (6) "Encountering the Literature," (7) "Responding to Literature," (8) "The Response Environment," (9) "Evaluating Students' Knowledge about and Understanding of Literature," and (10) "Supporting the Response to Literature Program." Completing all ten sections of the teacher survey takes 25–30 minutes. Teachers are asked to respond to most items according to the frequency with which they perform the kind of behavior indicated in the kind of class identified:

0 = Never—this is not something done in the type of class identified
1 = Infrequently—the behavior occurs no more than three or four times a year
2 = Sometimes—the behavior occurs at least five or six times a year but not as a regular practice
3 = Regularly—the behavior occurs throughout the year as a regular practice

Full instructions are provided for tabulating individual and group data.

Following is an example from each of the four Responding to Literature surveys:

Teacher (137 items):
As part of the way I evaluate students' abilities in literature, I
_____ 111. have students answer orally specific questions about details in the selections dealt with in class.

Student (52 items):
Before we read, our teacher would
_____ 11. tell us information about the author or the text.

Administrator (9 items):
_____ 6. I support teacher requests for aid to attend professional development sessions related to the responding to literature program.

Parent (9 items):
_____ 2. I talk to my child about the literature he or she reads, hears, or views in school.

Validity, Reliability, and Normative Data: Working documents "subject to further evaluation and revision based upon the outcomes of implementation activities" were reviewed. The author reports that

> responses from teachers completing initial drafts of the detailed teacher surveys have been very encouraging. Not only does it appear that completing the survey nurtures self-examination, but it also seems that the surveys provide for teachers very detailed catalogues of an extensive range of teaching activities organized into clear and understandable categories. In a very real sense, each survey is a framework for helping teachers understand how parts of various strategies and approaches interrelate and a basis for helping them integrate new ideas and techniques into a coherent instructional pattern.

The author acknowledges that everyone views events and processes from particular points of view. The compatibility of the content and response format of teacher and student response forms enables teachers to validate their own perceptions of what is happening in their classrooms. Although the teacher self-report instrument is the most comprehensive, the student, administrator, and parent surveys reflect important supportive behaviors and, indirectly at least, encourage including those potentially influential groups in school improvement efforts.

Ordering Information: ED 213 029

Category: Literature

Title: Social Relationships in Children's Stories (SRCS)

Author: Krishna Kumar

Age Range: Preschool and elementary children

Description of Instrument:

Purpose: To analyze the social relationships in stories for children. SRCS provides a compact design for multivariate content analysis of children's reading materials from a symbolic interactionist perspective. It was originally designed for a comparative study of the literary materials used for developing children's reading in Canada and India during grades 4, 5, and 6.

Date of Construction: 1980

Physical Description: SRCS is structured around the concept of "sociodrama"—the symbolic phases of action or "dramatic rehearsal" in which children participate while reading a story. It provides a framework of five "elements," and several categories within each element, as a means to analyze the roles symbolized in a children's story. The five elements are:

Agent (the central character)
Act (what took place in thought or deed)
Scene (the situation in which the act took place)
Agency (the means used in the performance of the act)
Purpose (the context or aim to which the act was addressed)

SCRS is presented in its entirety below:

Title ______________________ No. ____________

(Make as many choices as necessary under items 3.1, 4, and 5, but only one choice under items 1 and 3. Asterisk indicates need for detail.)

1. Agent (the character who performs the action depicted in a story or around whom the story revolves):

Child		Adult		Non-human	Group Agent
Male	Female	Male	Female		
1	2	3	4	5	6

1.2 Agent's background in terms of occupation and/or social status, if available (in the case of agents who are dependents of someone else, use the background of the person on whom they depend, e.g., parents in the case of child agent):

2. Act (summary of the plot-line, in terms of the agent's action, in one sentence):

A possible classification system for Acts:

1. Acts involving an attempt to survive under difficult circumstances
2. Acts involving encounter with evil characters
3. Acts of revenge
4. Acts which consist of doing good
5. Acts in which achievement of one's aim is involved
6. Acts performed out of deference to another
7. Acts in which making a choice leads to results
8. Acts involving routine life situations
9. Acts in which the agent responds to a supernatural/fantastic character/experience
10. Acts involving agent's response to coincidences

3. Scene:

Rural/Small Town	Urban	Unidentifiable
1	2	3

3.1 Space(s) used in the text:

Agent's Home	Socially Shared Spaces	Work Location	Nature	Institutional Spaces*	Another Person's House	Other*
1	2	3	4	5	6	7

4. Agency (means used in the performance of the act):

Personal Ability or Idea	Family	Peers	Community	Institutions	Non-human	Situation Improves Itself	Other*
1	2	3	4	5	6	7	8

5. Purpose (context in which the purpose of the act is established):

Self	Family	Peers	Community	Institutions	Work/ Money	Recreation	Moral/ Religious/ Altruistic Aims	Other*
1	2	3	4	5	6	7	8	9

6. Special comments ______________________________

Figure 1. Social Relationships in Children's Stories (SRCS)

Validity, Reliability, and Normative Data: SRCS is based on the "dramatistic pentad" proposed by Kenneth Burke (1945) and later recommended by Duncan (1968) as a means to study the structure of symbolic acts.

Agreement of 80.5 percent was reached between one coder and the developers of SRCS. For classification of acts alone (Item 2 in SRCS above) 64 percent agreement was recorded between the developer of SRCS and five judges.

Normative data on 79 stories in textbooks from India and 196 from Canada are available in Kumar (1982) below.

Ordering Information: Full instrument reproduced above.

Related Documents:

Burke, Kenneth. *A Grammar of Motives.* New York: Prentice-Hall, 1945.

Duncan, H. D. *Symbols in Society.* New York: Oxford University Press, 1968.

Kumar, Krishna. "Literature in the Reading Textbook: A Comparative Study from a Sociological Perspective." *Research in the Teaching of English* 16 (1982): 301-19.

Category: Literature

Title: Analyzing Characters in Literature (ACL)

Authors: National Assessment of Educational Progress Staff and Consultants

Age Range: Intermediate through postsecondary-adult

Description of Instrument:

Purpose: To evaluate students' written analyses of characters in prose fiction.

Date of Construction: 1979

Physical Description: After students read a story, they are first asked, "What kind of person is [character's name]? Describe [character's name] in a few words." Then they are asked, "What is it about the story that led you to describe [character's name] the way you did above? Write your answer below." ACL provides a scoring guide for students' answers to both questions. The guide provides three kinds of information about an essay: it tells whether the student identifies a character trait and substantiates it from the text, it gives the source of evidence from the text, and it gives the amount of evidence used for substantiation. ACL is presented in its entirety below. Examples included in some categories are from student responses to a particular story in the 1979–80 National Assessment of Reading/Literature. ACL is designed to be used with any story, however.

Scoring Guide Categories:

I. *First Categorization*—Identification and substantiation of character traits. This takes into account *both* questions.

1 = *Unable to identify character traits.* Respondents do not do the task. They refer to the text, but do not answer even the first question.

A. Only an opinion about the action of the character is offered, such as: "David shouldn't have left home."
B. Some material is quoted from the text with no clear identification of character (including quoting title).
C. The identification and substantiation of character seem unrelated to the text.
D. An observation about the story is made, for example: "The title is misleading."

2 = *Character trait identified without substantiation.* Respondents name something but cannot go on. They identify character trait(s) but do not substantiate the choice(s) with evidence from the text. Responses tend

to provide: 1) circular evidence, 2) a copy or close paraphrase of the text, 3) vague reasons, or 4) only a subjective reaction as substantiation.

3 = *Character trait identified and substantiated with minimal evidence.* Respondents identify character trait(s) and substantiate their choice(s) with *only one* reason or piece of evidence related to the text.

A. Reason can be directly related to the text, for example: "Nice, he wants to come home."
B. Reason can be inferred from the text.
C. Reason can be inaccurate, if it is related to the text, for example: "Smart, since he finally decided to go to college."
D. Reason can be based on personal experience that is related to the text.
E. Reason can be unusual, such as: "Sunburned from sitting out on the road," or "Lost, out in the road with no one around for miles."
F. Reason can refer to (but not retell) specific places in the text, for example: "The letter" is not specific enough. Also, referring to the place where specific adjectives were quoted from is merely a circular reason.

4 = *Character trait identified and substantiated.* Respondents identify character trait(s) and substantiate their choice(s) with *at least two* reasons or pieces of evidence related to the text. However, the evidence may be presented in an ambiguous fashion or be of the types described in 3C–3F. Reasons must be distinct—not instances of the same reason such as: "It had sad parts not any happy parts." This is a restatement of the same reason. Other instances of single reasons are when it takes two bits of information to make a single point, for example: (wanted to think things over) "He said he wanted to come home, but he didn't think he was ready for college."

5 = *Character trait identified and substantiated in a coherent fashion.* Respondents identify character trait(s) and substantiate their choice(s) with *at least two* reasons or pieces of evidence *clearly* related to the text—directly related or can be readily inferred. The reasons are presented logically and coherently.

Note: The following types of papers are classified as indicated and receive no further scoring:

0 = *No response.*
7 = *Illegible or illiterate.*
8 = *Totally off-task.*
9 = *"I don't know."*

II. *Second Categorization*—The source of the evidence. Code presence or absence for each of the following:

1 = *Content.* The evidence is based on the content of the text.
2 = *Form.* The evidence is based on the language, style, or construction of the text.
3 = *Subjective reactions.* These are responses that judge the worth of all or part of the text, such as: "It was interesting" or "It was monotonous." Personal opinions about the actions of the characters are stated, such as: "David should not have run away," or references to the moral of the story or general philosophical statements are made.

Note: Content and form can be present *only* if primary categorization is a "3" through "5"; subjective reactions can be present in papers categorized "2" through "5."

III. *Third Categorization*—A count of the number of reasons or pieces of evidence. Categorization for the count of details is as follows: (1), (2), (3), (4), (5), (6), (7 or more). Note: This count *only* applies to papers with primary categorization of "3" through "5"; subjective reactions should not be counted as reasons or evidence.

Validity, Reliability, and Normative Data: ACL would have content validity for a literature program seeking to encourage students to substantiate character interpretations with evidence from texts. Its construct validity is enhanced by the fact that it differentiates between the written responses of thirteen- and seventeen-year-olds.

After careful training of raters, the scoring constructor for National Assessment of Educational Progress was able to sustain rater agreement above 90 percent with ACL.

Since ACL was used in the 1979–80 National Reading/Literature Assessment, normative data are available for thirteen-year-olds and seventeen-year-olds. For example, 27 percent of the thirteen-year-olds and 41 percent of the seventeen-year-olds could identify and substantiate a character trait (fourth and fifth sections of the first categorization). For evidence, nearly all students in both age groups (73 percent of thirteen-year-olds, 84 percent of seventeen-year-olds) relied on content, rather than on form or subjective reaction. Further normative data are available in the National Assessment reports listed below.

Ordering Information: Full instrument reproduced above.

Related Documents:

National Assessment of Educational Progress. *Reading/Literature Released Exercise Set, 1979–80 Assessment.* Denver: National Assessment of Educational Progress, 1981. (ED 205 588)

———. *Reading, Thinking and Writing: Results from the 1979–80 National Assessment of Reading and Literature.* Denver: National Assessment of Educational Progress, 1981. (ED 209 641)

Category: Literature

Title: Analyzing Themes in Poetry (ATP)

Authors: National Assessment of Educational Progress Staff and Consultants

Age Range: Intermediate through postsecondary–adult

Description of Instrument:

Purpose: To evaluate students' written analyses of the ideas or themes in particular poems.

Date of Construction: 1979

Physical Description: Students are presented with a poem. They are then given these instructions: "Write an essay about an important idea or theme of the poem. In your essay tell how such things as the images, events, sounds, and structure contribute to this idea or theme." ATP is a guide to scoring students' essays. It is presented below in its entirety.

Scoring Guide Categories:

1 = *No analysis.* These responses only evaluate the poem or its features or make empty or glancing references to various features. Examples are: "All and all this poem was pretty and I enjoyed it," "It did have a lot of phrases that rhymed," "The poem presented poor images and events," "The structure was catchy," or "The sound is your singing a song."

Some category "1" responses do include brief allusions to the poem. However, these mentions of text are not considered synopsis. Also, wildly inaccurate interpretations of the poem and nonsensical responses should be placed in this category.

2 = *Synopsis.* These responses mainly retell or summarize the poem. Although some may include evaluations and empty or glancing references to other features, sometimes a brief synopsis can be embedded in an evaluation. If so, place it in category "2." The same is true of some references to images. When the meaning of an image is not given, but part of the poem is repeated, then the response can be placed in category "2." Also, responses that include glancing references to a number of features including *events* should be placed in this category. In summary, a "2" response at least retells, summarizes, or refers to particular parts of the poem. However, it does little else of substance in terms of analyzing the poem.

3 = *Theme.* These responses state an idea or theme of the poem, but do little of substance. They do not include synopsis or relevant discussions of other features. Some may include

evaluations, glancing references to features, or "philosophizing" about their theme that is not particularly relevant to the poem. In other words, some "3" responses may go off on tangents (initiated by the theme) which are not text based. A paper with no theme statement, but a substantive statement of one feature other than events, should also be placed in this category. For example: "An image is given." Some themes are: "It presented the idea of weakness in people," "The basic idea of the poem is how love hurts when misused or mishandled," "The theme is that you should not try to be somebody else," or "Always be your own person."

With a poem in particular, the difference between "synopsis" and "theme" is often a fine distinction. Yet the basic difference is whether or not the idea/message is stated as a generalization. Synopsis can involve hypotheses about meaning yet this is usually interpretation not generalization; for example: "I think it means he/she lost his/her identity." Also, some "3" responses elaborate their generalization to the point of directly referring to parts of the poem, such as: "By living by someone else's feelings, views, likes, and dislikes, talks and walks, you do not have an identity of your own." However, these should still be placed in category "3."

4 = *Minimal evidence of analysis.* Some of these responses state an idea or theme of the poem and relate events in the poem (plot summaries may be quite thin). References to specific parts of the text qualify as synopsis. Other responses placed in this category discuss *at least two* features; one can be events (synopsis), in a substantive manner—however, there is no statement of theme.

5 = *Evidence of analysis.* These responses state an idea or theme of the poem and include a substantive statement about *at least one* feature other than events (synopsis). For example, the structure might be discussed, such as: "The short lines and the choice of words suggest that this person is still not himself" or "The structure of this poem brought a melody of music, as though it were the lyrics of a soft-spoken song of tearful anger." These responses may also include synopsis or any of the other characteristics of papers placed in categories "1–4." However, the major drawback with category "5" papers is that they may be brief (include only one substantive statement), or, if they do include discussions of several features, the features explained or even elaborated do not all relate or refer back to the idea/theme proposed in the response. The relationship between the stated theme or idea of the poem and the discussion of the features is not explicit or even readily implicit.

6 = *Integrated analysis.* These responses state an idea/theme and discuss *at least two* features, one can be events (synopsis), in

> a substantive manner. They often have the components of "5" papers, yet the discussion of the features does relate to the proposed idea/theme. These are coherent, organized responses.
>
> Note: The following types of papers are classified as indicated and receive no further scoring:
>
> 0 = *No response.*
> 7 = *Illegible or illiterate.*
> 8 = *Totally off-task.*
> 9 = *"I don't know."*

Validity, Reliability, and Normative Data: ATP would have content validity for a literature program concerned with written analyses of poetry with an emphasis on justifying interpretations with evidence from the text.

After careful training of raters, the scoring contractor for National Assessment of Educational Progress was able to sustain rater agreement above 90 percent with ATP.

Since ATP was used in the 1979–80 National Reading/Literature Assessment, normative data are available for seventeen-year-olds. For example, only 5 percent of this age group was able to write responses scorable at the fifth or sixth level, responses that stated a theme and then went beyond mere synopsis of the text to provide evidence of the theme. Sixty percent of them could only summarize the poem or state a theme without evidence. Further normative data are available in the National Assessment reports listed below.

Ordering Information: Full instrument reproduced above.

Related Documents:

National Assessment of Educational Progress. *Reading/Literature Released Exercise Set, 1979–80 Assessment.* Denver: National Assessment of Educational Progress, 1981. (ED 205 588)

———. *Reading, Thinking and Writing: Results from the 1979–80 National Assessment of Reading and Literature.* Denver: National Assessment of Educational Progress, 1981. (ED 209 641)

Category: Literature

Title: Applying Criteria to Evaluate Literature (ACEL)

Authors: National Assessment of Educational Progress Staff and Consultants

Age Range: Junior high through postsecondary-adult

Description of Instrument:

Purpose: To assess the extent to which students can present criteria for their evaluations of particular poems or stories and elaborate the criteria with specific evidence from the text.

Date of Construction: 1979

Physical Description: ACEL is a descriptive scoring guide for students' open-ended written responses to the questions "Why is this a good (or bad) poem?" or "Why is this a good (or bad) story?" It identifies the number of criteria in the response, the extent to which each criterion has been elaborated upon, and the basis of evidence in each elaboration. The ACELs for analyzing and scoring responses to both poems and stories are reproduced below.

ACEL: For a Particular Poem

Scoring Guide Categories:

I. *First Categorization*—Presentation and elaboration of evidence.

1 = *No criteria or evidence given.* Respondent copies part of the text or gives a close paraphrase or circular response, for example: "It was good because it was good," "I liked it," "I didn't like it," or "I've heard it before." Nonsensical, or wildly inaccurate statements are given.

2 = *Gives a vague or unelaborated criterion.* A broad, sweeping generalization or personal assertion is made which does not necessarily have to restate the phrase "It was good/bad—." This response almost could have been given in absence of having heard or read the poem. It could apply to almost any poem. It was exciting, interesting, had a good plot, and so on (broad general adjectives).

3 = *Retells or gives summary or one vague criterion with synopsis as evidence.* The summary may refer to part or all of the poem; it may be cryptic, or lengthy and well written. This includes *any* citing of content of poem (as long as it is not basically copying).

4 = *Gives two or more unelaborated criteria.* Responses contain two or more generalizations or personal assertions. (These are longer "2s.")

5 = *Gives one criterion elaborated with evidence.* Respondent gives one criterion, generalization, or personal assertion that is supported with evidence other than retelling or plot summary; it may or may not be accompanied by unelaborated criteria. ("It was interesting because . . ."; respondent gives something other than plot summary.)

6 = *Gives two criteria elaborated with evidence.* Respondent gives two or more criteria, generalizations, or personal assertions *at least* two of which are supported with evidence other than retelling or plot summary; these may or may not be accompanied by unelaborated criteria. Note: Once a paper meets the criteria listed for a "4," "5," or "6" it does not matter if that response is also accompanied by plot summary.

Note: The following types of papers are classified as indicated and receive no further scoring:

0 = *No response.*
7 = *Illegible or illiterate.*
8 = *Totally off-task.*
9 = *"I don't know."*

II. *Second Categorization*—Basis of evidence. Code presence or absence for each of the following:

1 = *Content.* The evidence is based on the content of the text, for example: "It was about the crystal stair."

2 = *Form.* The evidence is based on the language, style, or construction of the text, for example: "There was so many misspelled words" or "Poems are supposed to rhyme."

3 = *Subjective reactions.* These are responses that judge the worth of all or part of the text, or give personal opinions about the actions of the characters, the believability of the plot, the moral of the poem, or the genre, such as: "I like poetry."

Note: Second categorization is *only* for papers with primary categorizations of "2" through "6."

ACEL: For a Particular Story

Scoring Guide Categories:

I. *First Categorization*—Presentation and elaboration of evidence.

1 = *No criteria or evidence given.* Respondent copies part of the text or gives a close paraphrase or circular response,

such as: "It was good because it was good," "I liked it," "I didn't like it," or "I've heard it before." Nonsensical or wildly inaccurate statements are given.

2 = *Gives a vague or unelaborated criterion.* A broad, sweeping generalization or personal assertion is made, which does not necessarily have to restate the phrase "It was good/bad—." This response almost could have been given in absence of having heard or read the story. It could apply to almost any story. It was exciting, interesting, had a good plot, and so on (broad general adjectives).

3 = *Retells or gives summary or one vague criterion with synopsis as evidence.* The summary may refer to part or all of the story; it may be cryptic, or lengthy and well written. This includes *any* citing of content of story (as long as it is not basically copying).

4 = *Gives two or more unelaborated criteria.* Responses contain two or more generalizations or personal assertions. (These are longer "2s.")

5 = *Gives one criterion elaborated with evidence.* Respondent gives one criterion, generalization, or personal assertion that is supported with evidence other than retelling or plot summary. It may or may not be accompanied by unelaborated criteria. ("It was interesting because . . ."; respondent gives something other than plot summary.)

6 = *Gives two criteria elaborated with evidence.* Respondent gives two or more criteria, generalizations, or personal assertions, *at least two* of which are supported with evidence other than retelling or plot summary. These may or may not be accompanied by unelaborated criteria. Note: Once a paper meets the criteria listed for a "4," "5," or "6" it does not matter if that response is also accompanied by plot summary.

Note: The following types of papers are classified as indicated and receive no further scoring:

0 = *No response.*
7 = *Illegible or illiterate.*
8 = *Totally off-task.*
9 = *"I don't know."*

II. *Second Categorization*—Basis of evidence. Code presence or absence for each of the following:

1 = *Content.* The evidence is based on the content of the text, for example: "Gives an idea of the old man's way of life."

2 = *Form.* The evidence is based on the language, style, or construction of the text, for example: "It didn't seem to have a beginning or an end."

3 = *Subjective reactions.* These are responses that judge the worth of all or part of the text, or are personal opinions about the actions of the characters, the believability of the plot, the moral of the story, or the genre, such as "I like fairy tales."

Note: Second categorization is only for papers with primary categorizations of "2" through "6."

Validity, Reliability, and Normative Data: Based on primary trait categories for persuasive writing (Lloyd-Jones 1977; Mullis 1980), ACEL was developed from a content analysis of responses from seventeen-year-old students. It would have content validity for a literature program which encouraged students to acquire a wide range of criteria for their evaluations of literature and the ability to support those criteria with either personal or textual evidence. ACEL differentiates across age levels.

After careful training of raters, the scoring contractor for National Assessment of Educational Progress was able to sustain rater agreement above 90 percent with ACEL.

Reports from the Third National Writing Assessment (1979–80) listed below provide complete normative data for seventeen-year-olds. For example, in responses to a story, only 10 percent could provide an acceptable response (at least two criteria elaborated with evidence). Thirty percent relied on summary or retelling to support just one criterion. Only 2 percent referred to the form of the text itself as evidence for any criteria.

Ordering Information: Full instrument reproduced above.

Related Documents:

Lloyd-Jones, R. "Primary Trait Scoring." In *Evaluating Writing: Describing, Measuring, Judging,* edited by C. R. Cooper and L. Odell. Urbana, Ill.: National Council of Teachers of English, 1977.

Mullis, I. V. S. *Using the Primary Trait System for Evaluating Writing.* Denver: National Assessment of Educational Progress, 1980.

National Assessment of Educational Progress. *Reading/Literature Released Exercise Set, 1979–80 Assessment.* Denver: National Assessment of Educational Progress, 1981. (ED 205 588)

———. *Reading, Thinking and Writing: Results from the 1979–80 National Assessment of Reading and Literature.* Denver: National Assessment of Educational Progress, 1981. (ED 209 641)

Category: Literature

Title: General Response to Literature (GRL)

Authors: National Assessment of Educational Progress Staff and Consultants

Age Range: Intermediate through postsecondary-adult

Description of Instrument:

Purpose: To describe students' written responses to works of literature.

Date of Construction: 1979

Physical Description: Students are asked to respond to literary works in any way that seems appropriate to them and given time to write out a response. GRL then permits a description or classification of the whole response and of individual statements within the response. Analysis of individual statements (propositions or T-units) reveals some of the personal, literary, and cognitive resources students bring to their responses to literature and their exploration and development of those responses in writing. GRL is presented below in its entirety. Though explorations of some descriptive categories contain references to a brief story and a poem used in the 1979-80 Reading/Literature Assessment, GRL is designed to describe comprehensively students' responses to any work of literature.

GRL: Story

Scoring Guide Categories:

EG = *Egocentric.* Responses are not text based, but are text relevant. Respondent writes a letter or story of his or her own or writes another story (or excerpts) that he or she has memorized. Other types of statements categorized here are: "I never read stories"; "I'm not good with stories"; or "I'm sorry to run out on you, I don't want to go to college either."

PR = *Personal—analytic.* Respondent gives personal reactions to content in an *analytic* sense—identification with characters, judgments about actions of characters and advice giving, observations about the way society should/does work. Respondent states, for example: "I might have done the same thing," "David shouldn't have left home," or "Hopefully his father will tie the cloth on the tree."

X = *Personal—global.* Respondent gives personal reactions to genre and content in a *global* sense. Examples would be statements of the following types: "I like stories about nature," "I wish I could write stories like this," or "This is not my kind of story."

EM = *Emotional.* Respondent attributes emotions or feelings of mood to the text or makes a direct statement of emotion. Examples would include: "The story was sad," "It's touching," "It had a funny feeling," "It was very dramatic," or "I felt sorry for the boy."

RT = *Retelling.* Respondent summarizes or retells the story (or parts of it). This summary can include statements referencing specific words or lines. (Disregard inaccuracies.)

IN = *Inferencing.* Respondent goes beyond the text and provides motivations for characters or develops action. Inferencing includes text-based hypotheses of what did happen or predictions about what will happen. For example: "David learned a lesson," "David's parents needed him to pay the bills," or "David feels that his father doesn't love him."

GN = *Generalization.* Respondent derives general meanings from the story, such as: "Go out and try new things," "It shows that people have feelings that can be hurt and people are the ones that hurt each other," or "Everyone knows you can't run away from your problems."

AN = *Analysis—superficial.* Respondent mentions *superficial* characteristics of the text. This includes concerns about format, for example: "It could have more details and not so many long words," "I didn't see any misspelled words," "It wasn't long enough," or "The author uses imaginative language."

Y = *Analysis—elaborated.* Respondent gives an *elaborated* or substantive discussion of any one of the following special features or literary devices: plot, characters, setting, images, sounds, etc. Included here are discussions of plot veracity and meaningfulness.

OW = *Other works—general.* Respondent classifies the work as to genre or type and compares the story to other types of works or art forms in *general,* for example: "It's not like a story I've seen before," "I think it's a good soap opera," or "It is like a myth."

Z = *Other works—specific.* Respondent compares the story to a *specific* work which is mentioned by title, such as: "The last paragraph reminds me of an old song, 'Tie a Yellow Ribbon 'Round the Old Oak Tree.'"

EV = *Evaluation.* Respondent judges the worth of the work. This judgment also includes such statements as: "It was stupid," "I didn't like it," "It doesn't make sense," "It is nicely written," "It was not exciting or sad," "It has no meaning," or "It is imaginative."

Note: In addition to the papers which are considered rateable (1 = rateable) and which are analyzed using the categories described above, some papers may not be considered rateable and these are placed in one of the following classifications:

0 = *No response.*
2 = *Nonrateable. Copies or uses circular statements.*
7 = *Illegible, illiterate.*
8 = *Totally off-task.*
9 = *"I don't know."*

GRL: Poem

Scoring Guide Categories:

EG = *Egocentric.* Responses are not text based, but are text relevant. Respondent writes a letter or poem of his or her own or writes another poem (or excerpts) that he or she has memorized. Other types of statements categorized here are: "I never read poetry," "I'm not good with poems," or "I love the beach—it helps me put my mind off things."

PR = *Personal—analytic.* Respondent gives personal reactions to content in an *analytic* sense—identification with characters, judgments about actions of characters and advice giving, observations about the way society should/does work. Respondent states, for example: "I might have felt the same thing," "It describes my feelings of moving to a new state," or "I feel that the poem is right by talking about the real problems of air pollution facing us."

X = *Personal—global.* Respondent gives personal reactions to genre and content in a *global* sense. Examples would be statements of the following type: "I like poems about nature," "I can relate to this poem," "I wish I could write poems like this," or "This is not my kind of poem."

EM = *Emotional.* Respondent attributes emotions or feelings of mood to the text or makes a direct statement of emotion. Examples would include: "The poem was sad," "It's touching," "It had a funny feeling," "It was very dramatic," "It gave a happy point of view," or "The ending makes you feel sorry for him."

RT = *Retelling.* Respondent summarizes or retells the poem or part(s) of it. This summary can include statements referencing specific words or lines. (Disregard inaccuracies.)

IN = *Inferencing.* Respondent goes beyond the text and provides motivations for characters or develops action. It includes text-based hypotheses of what did happen or predictions about what will happen, for example:

"The author is longing for the home he once had," or "A. E. Housman seems to be talking about a country that has been badly damaged or destroyed."

GN = *Generalization.* Respondent derives general meanings from the poem, such as: "Inside a man's heart live his fondest memories."

AN = *Analysis—superficial.* Respondent mentions *superficial* characteristics of the text. This includes concerns about format, for example: "The poem doesn't rhyme," "The poem seems more like a story," "It doesn't give the place and time," "The author uses imaginative language," or "There is a sense of lost beauty in the poem."

Y = *Analysis—elaborated.* Respondent gives an *elaborated* or substantive discussion of any one of the following special features or literary devices: plot, characters, setting, images, sounds, and so on. Included here are discussions of plot veracity and meaningfulness, such as: "Even though Housman wrote this poem in 1890, it is still pertinent and meaningful today," or "Each of us has memories of places and people we would like to relive, but which time will not allow us to."

OW = *Other works—general.* Respondent classifies the work as to genre or type and compares the poem to other types of works or art forms in *general,* such as: "It's not like a poem I've seen before," or "It is like a myth."

Z = *Other works—specific.* Respondent compares the poem to a *specific* work which is mentioned by title, such as: "The *Bible* describes heaven this way."

EV = *Evaluation.* Respondent judges the worth of the work. This judgment also includes such statements as: "It is stupid," "I didn't like it," "It doesn't make sense," "It is nicely written," or "It is imaginative."

Note: In addition to the papers which were considered rateable (1 = rateable) and which were analyzed using the categories described above, some papers were not considered rateable and these were placed in one of the following classifications:

0 = *No response.*
2 = *Nonrateable. Copies or uses circular statements.*
7 = *Illegible, illiterate.*
8 = *Totally off-task.*
9 = *"I don't know."*

Validity, Reliability, and Normative Data: GRL is a recent development in an established tradition of content analysis of written responses (Purves 1968; Odell and Cooper 1976). Developed from an analysis of responses of thirteen- and seventeen-year-olds, GRL would have validity for assessing changes in the emphasis or patterning of student responses as a result of instruction. For example, a goal of a secondary school literature course or program might be to decrease students' reliance on plot summary when they are invited to say anything they like about a short story they are reading for the first time. GRL is based on the assumption that many different kinds of responses may be appropriate for a given work of literature: personal, interpretive, analytic, comparative, or evaluative. Since each of these responses, or combinations of them, would be evaluated with different criteria, evaluators could develop a qualitative scale for each category in GRL. The construct validity of GRL is enhanced by the fact that it identifies differences in the response preferences of thirteen- and seventeen-year-olds.

After careful training of raters, the scoring contractor for National Assessment of Educational Progress was able to sustain rater agreement above 90 percent with GRL.

Since GRL was used in the 1979–80 National Reading/Literature Assessment, normative data are available for students aged thirteen and seventeen, both for frequency of appearance of GRL categories in any essay and for predominant response mode of each essay. For example, writing about a brief story, 57 percent of the thirteen-year-olds wrote essays which were "Personal—Analytic," whereas 67 percent of the seventeen-year-olds wrote essays in that mode. Ten percent of the younger group but only 6 percent of the older group wrote essays which were predominantly "Evaluative." Further normative data are available in the two National Assessment reports listed below.

Ordering Information: Full instrument reproduced above.

Related Documents:

National Assessment of Educational Progress. *Reading/Literature Released Exercise Set, 1979–80 Assessment.* Denver: National Assessment of Educational Progress, 1981. (ED 205 588)

———. *Reading, Thinking and Writing: Results from the 1979–80 National Assessment of Reading and Literature.* Denver: National Assessment of Educational Progress, 1981. (ED 209 641)

Odell, L., and C. R. Cooper. "Describing Responses to Works of Fiction." *Research in the Teaching of English* 10 (1976): 203–25. (EJ 151 216)

Purves, A. C., and Victoria Rippere. *Elements of Writing about a Literary Work: A Study of Response to Literature.* Urbana, Ill.: National Council of Teachers of English, 1968. (ED 018 431)

Category: Literature

Title: Qualities of Good Literature (QGL)

Authors: National Assessment of Educational Progress Staff and Consultants

Age Range: Intermediate through post secondary-adult

Description of Instrument:

Purpose: To assess the criteria students have for evaluating what makes a "good poem" or a "good story."

Date of Construction: 1979

Physical Description: QGL is a descriptive scoring guide for students' open-ended written responses to the questions "What makes a good poem?" or "What makes a good story?" QGL provides two kinds of information about students' responses: the specific criteria or qualities students mention about stories or poems and the levels of the criteria, whether concerned with features of the work as a whole, the ways texts function, or the meanings of texts. The QGLs for both poems and stories are reproduced below.

QGL: Good Poem

Scoring Guide Categories:

I. *First Categorization*—Description of type of qualities listed.

0 = *No response.*

1 = *Relationship between form and content.* Respondents may state the rhyme pattern is compatible with the topic, for example.

2 = *Content.* Respondents may refer to one or more of the following types of content: mystery, western, fantasy, adventure, danger, action, humor, suspense, romance, drama, any reference to theme, or excitement.

3 = *Form.* Respondents mention some aspect of form, such as: good rhyme, high point, strong words, vivid language, suspenseful beginning, length, style, construction of the text, or happy ending.

4 = *Subjective reaction.* Respondents give a statement to the effect that a poem should evoke a subjective reaction of one of the following types: sensible, interesting, intelligent, funny, imaginative, dramatic, suspenseful, or adventurous.

5 = *Unelaborated features of genre.* Respondents refer to one of the following characteristics: plot, character, setting—with or without redundant "good."

6 = *Naming of a specific poem or poet.* Respondents list a particular poem or poet.

7 = *Undetermined or circular.* These are responses where you cannot determine whether the quality is one of content or form, or the answer is circular, for example: poet, good poet, good literature, good writing, good words, the title, the ending, language, or good subject.

8 = *References to format.* Respondents list some quality related to format, for example: neatness, commas, quotation, indentations, capital letters, summary, controlling idea, or has a title.

9 = *Other.* Responses are totally off-task, illegible, illiterate, "I don't know," or nonsensical.

II. *Second Categorization*—Level of the qualities listed.

1 = *Identifies characteristics of work as a whole.* Responses refer to sex, violence, human adventure, catchy title, dialogue, good rhythm, plot, or setting.

2 = *Analyzes the way the text works.* Respondent states ideas, such as: use of foreshadowing or irony, the tension rises, enthusiastic words are used, or surprise ending is effective.

3 = *Makes statements about the meaning or theme.* Respondent makes statements, such as: "It makes me think"; "It has a good moral"; "It expresses your feelings"; or "It has meaning, meaningfulness."

Note: This categorization *only* applies to papers rated "1" through "5" for the first categorization.

QGL: Good Story

Scoring Guide Categories:

I. *First Categorization*—Description of type of qualities listed.

0 = *No response.*

1 = *Relationship between form and content.* Respondents may state that the dialogue is compatible with the topic, for example.

2 = *Content.* Respondents may refer to one or more of the following types of content: mystery, western, fantasy, adventure, danger, action, humor, suspense, romance, drama, any reference to theme, or excitement.

3 = *Form.* Respondents mention some aspect of form, such as: high point, strong words, vivid language, suspenseful beginning, length, style, construction of the text, or happy ending.

4 = *Subjective reaction.* Respondents give a statement to the effect that a good story should evoke a subjective

reaction of one of the following types: sensible, interesting, intelligent, funny, imaginative, dramatic, suspenseful, or adventurous.

5 = *Unelaborated features of genre.* Respondents refer to one of the following characteristics: plot, character, setting—with or without redundant "good."

6 = *Naming of a specific story or author.* Respondents list a particular title or author.

7 = *Undetermined or circular.* These are responses where you cannot determine whether the quality is one of content or form, or the answer is circular, for example: the writer, author, good author, good literature, good writing, good words, the title, the ending, language, or good subject.

8 = *References to format.* Respondents list some quality related to format, for example: neatness, commas, quotation, indentations, capital letters, summary, controlling idea, or has a title.

9 = *Other.* Responses are totally off-task, illegible, illiterate, "I don't know," or nonsensical.

II. *Second Categorization*—Level of the qualities listed.

1 = *Identifies characteristics of work as a whole.* Responses refer to sex, violence, human adventure, catchy title, dialogue, plot, or setting.

2 = *Analyzes the way the text works.* Respondent states ideas, such as: use of foreshadowing or irony, sentences are to the point, the tension rises, enthusiastic words are used, surprise ending is effective.

3 = *Makes statements about the meaning or theme.* Respondent makes statements, such as: "It makes me think"; "It has a good moral"; "It expresses your feelings"; or "It has meaning, meaningfulness."

Note: This categorization *only* applies to papers rated "1" through "5" for the first categorization.

Validity, Reliability, and Normative Data: QGL was developed from a content analysis of responses from students ages nine, thirteen, and seventeen. QGL would have content validity for a literature program in which students were expected to develop a concept of "good" literature and express the concept in writing with specific criteria. QGL differentiates across the age range from nine to seventeen.

After careful training of raters, the scoring contractor for National Assessment of Educational Progress was able to sustain rater agreements above 90 percent with QGL.

Since QGL was administered in the Third National Writing Assessment (1979–80) to a stratified random sample of American students, it comes with a large amount of normative data for the three age groups nine, thirteen, and seventeen. For example, 83 percent of the seventeen-year-olds could give three criteria for a good story, but only 40 percent of the nine-year-olds could give three criteria. Older students relied more on criteria that were based on content and literary features, while younger students' criteria tended to be undeterminable or circular. Further normative data are in the two reports listed under "Related Documents" below.

Ordering Information: Full instrument reproduced above.

Related Documents:

National Assessment of Educational Progress. *Reading/Literature Released Exercise Set, 1979–80 Assessment.* Denver: National Assessment of Educational Progress, 1981. (ED 205 588)

———. *Reading, Thinking and Writing: Results from the 1979–80 National Assessment of Reading and Literature.* Denver: National Assessment of Educational Progress, 1981. (ED 209 641)

Category: Literature

Title: Story Preference Inventory (SPI)

Author: Philip E. Swayne

Age Range: Primary through intermediate

Description of Instrument:

Purpose: To assess children's preferences for story settings

Date of Construction: 1975

Physical Description: SPI is a forced-choice instrument upon which children indicate their preference for a story setting in each of thirty-six paired pictures. Each pair of pictures contains two contrasting story settings. There are twelve pairs of pictures in each of the following three categories:

1. Fantasy vs. realism. Example: A picture of a physician examining a child with a stethoscope is paired with a contrasting picture in which an adult bird is examining a young bird with a stethoscope.
2. The past vs. the contemporary. Example: A picture of a boy doing his homework on the floor in a rustic log cabin by the light of a fireplace is paired with a contrasting picture in which a boy in modern dress and surroundings is doing his homework at a desk.
3. The geographically remote vs. the near-at-hand. Example: A picture of a Japanese girl in a Japanese setting playing a koto (Japanese zither) is paired with a contrasting picture of a girl in the United States in a characteristic setting playing a guitar. The Inventory represents foreign settings that persist as parts of a nation or culture.

Each of the thirty-six pairs of pictures appears on a separate page. The children are instructed to look at the two contrasting pictures on each page and to select that picture which they would rather read a story about (or, in the case of nonreaders, which they would rather *hear* a story about). The sequence of the items on a page and the sequence of the thirty-six pages were randomly determined.

The inventory yields three scores, one for each of the categories described above. A score of one is given for each response that shows a preference for fantasy, for the past, or for the geographically remote. Preference for a contrasting setting is scored zero. Thus, for any of the three categories, the range of possible scores is 0–12.

Validity, Reliability, and Normative Data: The author showed the picture pairs individually to ten second-grade pupils in a Philadelphia public school whom the principal selected as representative of the ability of the school's second graders. The author asked each pupil to tell the difference between the contrasting pairs of pictures. Responses were accurate with the following exceptions: Two pupils were unable to differentiate two pairs, and a third pupil was unable to differentiate two other pairs. The pictures in those four pairs were redrawn, and the validating process repeated with a second group of ten pupils from the same second-grade class, similarly judged by the principal to be representative. On this occasion, all differentiations were successful.

An additional test of validity was undertaken by administering the inventory to twelve randomly selected pupils from all second-grade sections of the same school. The author subsequently interviewed these pupils individually, showing pupils their inventory and asking why they had chosen as they did. The pupils' explanations revealed an understanding in each instance of the intended contrast.

The inventory was then administered to all pupils ($n = 49$) in grades 2, 4, and 6 of a Philadelphia suburban school. In scoring, the inventory was randomly divided into two halves, selecting eighteen of thirty-six items—six items from each category. The scores of the halves were correlated and the correlations corrected by the Spearman-Brown formula, yielding reliability coefficients as follows:

Fantasy vs. realism	.96
The past vs. the contemporary	.89
The geographically remote vs. the near-at-hand	.86

No normative data are available.

Ordering Information: ED 236 639

Related Documents:

Carter, S. M. "Interests and Reading." *Journal of Research and Development in Education* 11 (1978): 11–68.

Pieronek, F. T. "Do Basal Readers Reflect the Interests of Intermediate Students?" *Reading Teacher* 33 (1980): 408–12. (EJ 225 173)

Category: Literature

Title: Criteria for Evaluating Picture Story Books (CEPSB)

Author: Jeane Sword

Age Range: Postsecondary–adult

Description of Instrument:

Purpose: To provide teachers in early childhood education, as well as other interested educators, with criteria for evaluating the quality of plot, characterization, and style in picture storybooks. This instrument makes no attempt to evaluate the illustrations of the picture storybook.

Date of Construction: 1980

Physical Description: A set of eight criteria is used to evaluate the literary elements of plot, characterization, and style in picture storybooks. Each criterion is followed by statements that illustrate different ways an author might fulfill a criterion. The evaluator is to indicate the degree to which the author has successfully achieved the standard by placing a sign in front of each: a plus sign (+) indicates excellent; a check mark (✓) indicates satisfactory; and a minus sign (−) indicates minimal achievement of the criterion. The evaluator then rates each criterion on a 1- to 5-point scale, with 1 point indicating a low rating and 5 points a high rating. Below is one criterion from the eight in CEPSB:

1. How well did the author achieve plot unification?

   Below are listed several ways in which to achieve plot unification. Indicate the degree to which this particular author successfully achieves that quality.

   _____ Beginning of story establishes conflict.
   _____ Middle of story presents plausible obstacles.
   _____ Middle of story presents rising action.
   _____ End of middle section has an identifiable climax.
   _____ Ending is brief.
   _____ Ending is satisfying.

   Based on your responses to the above question, rate this book on plot unification on a 1–5 scale.

1 low	2	3	4	5 high

Validity, Reliability, and Normative Data: CEPSB was content validated by the ratings of three expert judges. These experts were university faculty members who taught children's literature and who had published extensively in the field. The judges evaluated each criterion and recommended one of three courses: (1) to retain the criterion, (2) to delete the criterion, or (3) to change the criterion (sheets were attached for recommendations). All three judges recommended retention of the eight criteria in CEPSB.

Three early childhood and elementary education teachers and seven university students of children's literature used the instrument to rate a set of ten books; these ratings were then correlated. The interrater reliability of the instrument was statistically significant as shown by Kendall Coefficient of Concordance, $W(10,10) = .80, p < .01$. The concurrent validity was also assessed by correlating the median student ratings for each book with this author's ratings as a criterion. The concurrent validity as indexed by Spearman Correlation Coefficient was $r_s(10) = .67, p < .05$.

Ordering Information: ED 236 658

Related Document:

Sword, J. M. "Factors Related to Kindergarten Teachers' Book Selection." Ph.D. diss., University of Illinois, 1979. *Dissertation Abstracts International* 40: (1980) 5312-A.

Oral Language

Category: Oral Language

Title: Oral Composing. Communication Skills. Pennsylvania Comprehensive Reading/Communication Arts Plan (PCRP) Assessment Survey III

Author: Stephen M. Koziol, Jr. for the Pennsylvania Department of Education

Age Range: Adult

Description of Instrument:

Purpose: To serve as a vehicle for individual elementary and secondary school teachers to examine and reflect upon their own curriculum designs and instructional practices in teaching oral composing and to serve as a tool for schools involved in evaluating their instructional programs and planning long-range improvements.

Date of Construction: 1982

Physical Description: Oral Composing is one of five survey packets developed to gather accurate, detailed, comprehensive, and descriptive data about instruction in communication skills in a single classroom or course. Packets on the subjects of literature, written composing, oral composing, language proficiency development, and sustained silent reading were designed for teachers engaged directly in communication skills instruction as well as for their students, the parents of their students, and their school administrators. Each of the five surveys has four separate forms (one for teachers, one for students, one for administrators, and one for parents) and each addresses two layers of specificity: responses to aspects of the overall communication skill program and responses to a teacher's practices and policies within a single kind of class situation.

The teacher forms are divided into sections related to distinct components of instruction. The Oral Composing form has eight main sections: (1) "Background Information," (2) "Goals for Oral Composing," (3) "Pre-Composing Activities," (4) "Types of Assignments," (5) "Audiences," (6) "Oral Composing Environment," (7) "Response and Evaluation," and (8) "Supporting Work." Teachers are asked to respond to most items according to the frequency with which they perform the behavior indicated in the kind of class identified:

0 = Never—this is not something done in the type of class identified
1 = Infrequently—the behavior occurs no more than three or four times a year

2 = Sometimes—the behavior occurs at least five or six times a year but not as a regular practice

3 = Regularly—the behavior occurs throughout the year as a regular practice

Full instructions are provided for tabulating individual and group data.

Following is an example from each of the four Oral Composing surveys:

Teacher (110 items):
In preparing students for oral composing assignments, I
_____ 15. lead "brainstorming" sessions on specific topics or assignments.

Student (50 items):
_____ 9. The teacher would have us listen to examples or models.

Administrator (12 items):
_____ 6. I inform students and staff about oral composing events in the community or on the radio.

Parent (9 items):
_____ 1. I read things orally to or with my child.

Validity, Reliability, and Normative Data: Working documents "subject to further evaluation and revision based upon the outcomes of implementation activities" were reviewed. The author reports that

> responses from teachers completing initial drafts of the detailed teacher surveys have been very encouraging. Not only does it appear that completing the survey nurtures self-examination, but it also seems that the surveys provide for teachers very detailed catalogues of an extensive range of teaching activities organized into clear and understandable categories. In a very real sense, each survey is a framework for helping teachers understand how parts of various strategies and approaches interrelate and a basis for helping them integrate new ideas and techniques into a coherent instructional pattern.

The author acknowledges that everyone views events and processes from particular points of view. The compatability of the content and response format of teacher and student response forms enables teachers to validate their own perceptions of what is happening in their classrooms. Although the teacher self-report instrument is the most comprehensive, the student, administrator, and parent surveys reflect important supportive behaviors and, indirectly at least, encourage including those potentially influential groups in school improvement efforts.

Ordering Information: ED 213 031

Category: Oral Language

Title: Communication Competency Assessment Instrument (CCAI)

Author: Rebecca B. Rubin

Age Range: College

Description of Instrument:

Purpose: Designed to assess the student's ability to function in college contexts, the CCAI is a direct and comprehensive measure of college-level speaking and listening skills.

Date of Construction: 1981

Physical Description: The report of the Speech Communication Association's Task Force on Minimal Speaking/Listening Competencies formed the structure of the CCAI. This report identified four major competency areas: Communication Codes (the ability to use and understand spoken English and nonverbal signs), Oral Message Evaluation (the ability to use appraisal standards to judge oral messages and their effects), Basic Speech Communication Skills (the ability to select and arrange message elements to produce spoken messages), and Human Relations (the ability to maintain interpersonal relationships). The report further divides these four areas into nineteen specific competencies, with examples of application from three contexts: occupational, citizenship, and maintenance. For the CCAI a fourth context—educational—was created, along with three application examples for each of the nineteen competencies. Fifty-seven possible assessments were thus made about a student's ability to function in specific educational environments: in class, and with instructors, fellow students, and academic advisors. In this fifty-seven-item, hour-long version of the CCAI, the student is first asked to present a three-minute, extemporaneous, persuasive talk on a topic of interest, during which the student's speaking ability is judged on factors such as volume, rate, clarity, and gestures. Next, the student views a videotaped class lecture, is questioned immediately about the lecture, and is asked to respond to statements about his or her experiences in an educational environment.

The first version of the CCAI was critically reviewed and refined to reduce testing time. The result is a 30-minute-per-student, nineteen-item short form. Students are rated on a 5-point scale for: (1) pronunciation, (2) facial expression/tone of voice, (3) speech clarity, (4) persuasive/informative distinction, (5) clarity of ideas, (6) expression and defense of a point of view, (7) recognition of nonunderstanding, (8) fact/opinion distinction, (9) listening (understanding of suggestions), (10) identification of main ideas, (11) summarization, (12) social ritual, (13) question asking,

(14) question answering, (15) expression of feelings, (16) organization of ideas, (17) direction giving, (18) description of another's view, (19) description of a difference in opinion. Those receiving 57 total points (average grade of 3 for nineteen items) are declared "competent"; those above 48 points (2.5 average on a 5-point scale), "in need of remedial self-paced work"; and those below 48 points, "in need of formal training in a remedial course."

Validity, Reliability, and Normative Data: The CCAI is predicated on the Speech Communication Association's (SCA) "Minimal Competencies in Speaking and Listening for High School Graduates," which identifies the four competency areas of communication codes, oral message evaluation, basic speech communication skills, and human relations. It also follows the SCA's "Criteria for Evaluating Instruments and Procedures for Assessing Speaking and Listening."

The CCAI has undergone extensive testing at the University of Wisconsin–Parkside. Content validity of the long form was achieved by presenting five communication faculty members with the nineteen competencies and fifty-seven assessment items arranged in random order and asking them to place each assessment item into one of the nineteen competency categories. Five questions that initially failed to meet an established 80 percent agreement standard were subsequently rewritten and correctly categorized by all evaluators. A rating booklet with five levels of proficiency for each of the assessments was constructed, evaluated, and revised. Four faculty members trained as judges used the booklet to evaluate three students who had been videotaped while taking the CCAI. A mean correlation of .83 was attained. Seventy-seven students were subsequently assessed with this test version.

In an effort to develop a short form, coefficient alpha analysis was performed on the fifty-seven-item CCAI. An overall alpha of .33 was achieved. The least consistent items were eliminated; that is, for each of the nineteen competency areas the most valid item was determined. The coefficient alpha for the nineteen-item form was .79.

Data are available relating performance on the CCAI to sex, ethnicity, academic major, credits completed, age, grade point average, number of communication or speech courses taken, and past speaking experience.

Ordering Information: ED 210 748

Category: Reading

Title: Assessment of Instructional Terms (AIT)

Authors: Idrenne Alparaque
William T. Fagan

Age Range: Primary grades

Description of Instrument:

Purpose: The Assessment of Instructional Terms was designed to assess first-graders' knowledge of commonly used instructional terms in a particular language arts program.

Date of Construction: 1983

Physical Description: The AIT consists of two parts—a verbal part and a situational part. In each part, the same twelve commonly used instructional terms are assessed. These are: *word, begin, letter, name, makes sense, beginning sounds, print, tract, capital letter, rhyme, period, stands for*. In the situational task, the terms are assessed in a worksheet format. The directions are phrased so that the term focused on is the only technical term used.

The verbal part consists of three questions for each term. The first is a general question simply asking the child "What does _____ mean to you?" The second and third questions seek a function and exemplar of the word, respectively. The first question is designed so that the responses can be analyzed descriptively. The responses to the second and third questions may be analyzed statistically and compared to the response of the situational tasks that require the child to find an example in a functional situation.

A sample item from the verbal part of the AIT for the term *capital letter* is given below:

(Verbal Task)
a. What does *capital letter* mean to you?
b. Where would you use a *capital letter*?
c. Can you show me a *capital letter*?

Validity, Reliability, and Normative Data: Content validity may be claimed for the AIT from a number of perspectives. (1) The terms assessed were those that commonly occurred in a language arts program at the primary level (*Starting Points in Language Arts,* Level I, Ginn). (2) Six criteria or guidelines for each task (verbal and situational) were developed and followed in the construction of the instrument. (3) Three graduate students evaluated the tasks in terms of the guidelines developed. (4) Two pilot studies were conducted in order to finalize the task questions.

Construct validity may be claimed for the instrument in the sense that Fisher (1981) served as a theoretical framework for elucidating the steps necessary to complete the task and assisted with the formation of the guidelines. Finally, ecological validity is based on the results of the study that showed that the instrument was able to differentiate the knowledge of a high- and low-reader group at the grade one level.

Interrater reliability was established on 10 percent of the protocols of the situational task and questions two and three of the verbal part (used for statistical anaylsis) and on 20 percent of the protocols for question one of the verbal part (analyzed descriptively). The level of agreement between two independent raters for all analyses was 100 percent.

Ordering Information: ED 236 642

Related Documents:

Alparaque, I. "The Young Child's Understanding of Commonly Used Instructional Terms in Language Arts." Master's thesis, University of Alberta, Edmonton, Alberta, Canada, 1983.

Fisher, D. L. "Functional Literacy Test: A Model of Question-answering and Errors." *Reading Research Quarterly* 15 (1981): 418–47. (EJ 245 664)

Category: Reading

Title: A Test of Staging Effects on Recall from Prose

Author: Paul Clements

Age Range: High school, postsecondary

Description of Instrument:

Purpose: To determine whether staging of information (its degree of prominence—high or low) in a passage would influence a reader's recall.

Date of Construction: 1976

Physical Description: The task consisted of two pairs of passages. The members of each pair had identical content but different staging. Passages one and two were about the sighting of a rare bird, while the second pair (three and four) dealt with the topic of psychosurgery. In the first two passages, there were four information chunks of interest (labeled A, B, C, D) and in the second pair there were two such chunks (labeled E, F). Subjects were asked to read the passage twice at their normal speed, and then to write all that they could remember of the passage. A sample passage, followed by its "staging structure," is given below:

> Grapes for wine-making require a temperate climate but table grapes flourish in warmer climates. California provides the climatic range needed for both kinds of grapes. Californian vineyards produce high quality wines but some regions of California are too warm for wine-producing grapes. Warmer climates produce grapes with a high sugar content. Such grapes are good to eat but are not ideal for wine-making.

Signal	Marker	Topic	Comment
Start		Grapes for wine-making	require a temperate climate
Coord	but	table grapes	flourish in warmer climates
New		*California	provides the climatic range needed for both kinds of grapes
New		**Californian vineyards	produce high quality wine
Coord	but	**some regions of California	are too warm for wine-producing grapes
Old		Warmer climates	produce grapes with a high sugar content
Old		Such grapes	are good to eat
Coord	but	(such grapes)	are not ideal for wine-making

There are four comments that should be made about this example.

a. The "signal" column shows the basis on which the hierarchical level is assigned. Level is shown by the number of stars preceding the topic. No stars means that the topic (and its comment) are at the highest level in the structure, one star means one level down, etc.

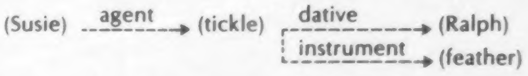
b. The first topic is at the highest level for definition. If the passage had had a title, the title would have been the first topic.

c. Towards the end of the passage, "warmer climate" is "old" because it occurred in the comment of the second topic. "Such grapes" is "old" because it is referentially the same as "grapes with a high sugar content" from the previous comment.

d. The "marker" column shows when explicit linguistic signals are used.

Validity, Reliability, and Normative Data: Construct validity was established by arguing that staging is an aspect of discourse processing and is relevant to understanding the cognitive processes that occur during reading. The author's discourse production model (Clements 1976, 44) was based largely on the work of Grimes. It is further stated that within a discourse processing model, staging possesses explanatory value in that it provides a basis on which syntactic choices can be made that would otherwise be unmotivated.

In order to devise the passages for the task, it was necessary to set up rules (1) for mapping the staging structure of a passage and (2) for representing the semantic content for scoring. Eight rules were devised to accomplish the first step (26–29). Fredericksen's (1972) system was used to represent the semantic content of the passages. Charts were prepared that depicted the semantic structures of the chunks to be scored. A brief example is provided below to illustrate the method. In Fredericksen's system, the sentence "Susie tickled Ralph with a feather" would be represented as:



Thus, the system represents propositions as slots connected by labeled relations. In this example, then, Susie is the agent of the action "tickle," Ralph is the recipient ("dative"), and a feather is the instrument. There are seven score points in this example—four slots and three relations. If a subject's recall said "Someone tickled Ralph with a feather," then the first slot would not be scored but the remaining six data points would be, because the recall says that some agent tickled Ralph with a feather. Given the recall "Susie tickled Ralph for fun," then the first five data points would be scored, but not the last two because no mention is made of a feather or of any instrument.

A separate semantic structure chart was prepared for every subject recall. Those elements that were present in a recall were marked on the subject's chart for that passage. The number of data points present in each information chunk was then tallied for each subject. Thus, each subject was given a score for each chunk he or she recalled.

In order to obtain reliability, a sample of recalls containing eighteen information chunks was rescored by another investigator trained in the use of Fredericksen's method. This investigator was unaware of the direction of the hypotheses and of any of the experimental details. There are two measures of agreement that may be considered. One is the number of data points on which both scorers agreed. The other is the correlation between raters for scores on each chunk. It is the latter measure that was crucial for the study, but the former more detailed comparison may be of interest to investigators who plan more fine-grained semantic analyses than the between-chunk comparisons used in this study.

In the eighteen information chunks which were rescored, there were 762 data points. There was agreement on 721 of these and disagreement on 41. Thus there was agreement on 94.6 percent of the data points, which yielded a phi correlation of .87, based on a two-way contingency table showing the number of points scored as present or absent from recalls by each rater. The between-scorer product-moment correlation for chunk scores was +.95. The results show that, despite occasional disagreement over detail, the scores determined for chunks of information in this study are highly reliable.

In general, results from a sample of high school and college subjects showed that there was more recall for information when it was staged higher in the passage.

Ordering Information: University Microfilms, order no. 76-14, 848.

Related Documents:

Clements, Paul. "The Effects of Staging on Recall from Prose." Ph.D. diss., Cornell University, 1976.

Fredericksen, C. H. "Representing Logical and Semantic Structure of Knowledge Acquired from Discourse." *Cognitive Psychology* 7 (1972): 371–458.

Grimes, J. H. *The Thread of Discourse.* The Hague: Mouton, no date available.

Category: Reading

Title: The Interaction of Picture and Print in Reading Instruction

Author: Susan Dalfen Denburg

Age Range: Primary

Description of Instrument:

Purpose: To consider whether pictures might facilitate word identification and word learning and to determine the most appropriate design of pictures to aid in independent reading.

Date of Construction: 1975

Physical Description: Pictures have been criticized for attracting attention away from the text and for presenting ambiguous and/or misleading information; they are also praised for arousing a child's attention and for motivating him or her to read. However, the role of pictures in reading has rarely been delineated, nor have pictures been designed to optimize that role. To fulfill this need, this instrument consisted of twenty-four sentences and accompanying pictures that completely or partially represented the noun information of the sentence.

The sentences were pasted onto the bottom of an 8½-by-10-inch piece of white bristol board, and a glossy photograph of the matching picture (approximately 5-by-6-inch) was pasted onto the top half of the page. A test booklet consisted of twenty-four such pages—six with a sentence alone, six with a sentence plus a full picture, six with a sentence plus a partial picture (object noun deleted), and six with a sentence plus a partial picture (subject noun deleted).

The general directions were:

> Here is a book that we're going to read. On every page you'll see a short sentence that tells a little story. Sometimes there will be a picture that goes with the words in the sentence and sometimes the sentence will be all alone on the page. I want you to look at the words in the sentence carefully and after you read them to yourself, read them to me out loud. Many of the words will be hard for you to read, but I want you to try your best to read them to me.
>
> If there is a picture on the page, use it to help you figure out the words in the sentence. Sometimes the picture will go with all the words in the sentence and sometimes the picture will have a missing part and show you only part of the sentence. You should use the picture to help you with the words but remember that I want you to read me the words in the sentence, so always look at the sentence when you read.

> The description of an example is as follows: for the sentence "The donkey pulls the wagon," the full picture portrays a donkey at a slight angle with a wagon attached to his sides and trailing behind him. Deleting the donkey leaves the wagon attached to something (obviously missing) and deleting the wagon leaves a donkey with some sticks attached to his sides and pointing behind him.

Substitutions for the target words (pictured) were categorized according to their degree of "graphic proximity" (Goodman and Burke 1973) and contextual congruency. Graphic proximity refers to the degree of graphemic similarity between observed and exported (correct) response.

Validity, Reliability, and Normative Data: The identification of a word as a construct during fluent reading was considered to result from an interplay of information from various sources, including graphemic information available in the printed word and contextual information which encompasses knowledge of syntactic and semantic constraints of the language. Further, a trade-off is postulated between visual and nonvisual information so that the more of the latter that the reader can contribute, the less of the former he or she need sample in order to identify a word correctly. This trade-off is seen as partial rather than complete since information from the various sources is pieced together or integrated and works cooperatively in word identification.

Within the framework of this construct, a number of principles guided the construction of the instrument:

1. The combination of picture and text should provide a "communication package" which surpasses the interest or communication potential of either alone.
2. The word cued by the picture should not be so predictable that it obviates the need to examine the printed word.
3. Conversely, the relationship must not be so unpredictable as to be bizarre.
4. The words pictured must be in the child's listening or spoken vocabulary.
5. The sentences used should be based on the degree of syntactic awareness that first graders have been shown to possess.

Sentence patterns chosen for the instrument were based on data obtained by Menyuk (1963) while words chosen were those likely to be part of the vocabulary of first graders according to the Thorndike and Lorge (1944) word count and the Rinsland (1945) oral vocabulary list. The final selection of words and pictures was determined through pilot studies, and the judgments of a sample of adults.

The interrater reliability of two adults for the graphic proximity scores was 0.99.

Data from forty-eight first graders indicate that the presence of a picture has a strong facilitative effect on word identification in context and a smaller though significant facilitative effect on word learning, as measured by a follow-up word list task. The availability of pictorial information seems to provide the children with an opportunity to use whatever limited graphemic skill they possess and to integrate graphemic with contextual information in a manner reminiscent of the skilled reader. (Refer to abstracted report in *Reading Research Quarterly* for particular data regarding degree of readers' graphemic knowledge and optimal amount of pictorial information to achieve maximal gains).

Ordering Information: ED 236 662

Related Documents:

Denburg, Susan Dalfen. "The Interaction of Picture and Print in Reading Instruction" (abstracted report). *Reading Research Quarterly* 12 (1976–77): 176–89. (EJ 153 486)

Denburg, Susan M. "The Interaction of Picture and Print in Reading Instruction." Ph.D. diss., University of Toronto, 1975.

Goodman, K. S., and C. L. Burke. *Appendix for Theoretically Based Studies of Patterns of Miscues in Oral Reading Performance.* Final Report. Office of Education, Department of Health, Education, and Welfare, April, 1973.

Menyuk, P. "Syntactic Structures in the Language of Children." *Child Development* 34 (1963): 407–22.

Rinsland, H. D. *A Basic Vocabulary of Elementary School Children.* New York: Macmillan, 1945.

Thorndike, E. L., and I. Lorge. *The Teacher's Wordbook of 30,000 Words.* New York: Teachers College, Columbia University, 1944.

Category: Reading

Title: The Proposition Inventory

Authors: Gerald Duffy
Linda Anderson

Age Range: Adult

Description of Instrument:

Purpose: To characterize how teachers use reading theories, models, and other conceptions as they plan and carry out reading instruction.

Date of Construction: 1982

Physical Description: The Proposition Inventory is a forty-five-item questionnaire with 5-point Likert scoring. Propositions were derived from searches of standard reading methods texts reflecting various theoretical perceptions of the field, generally categorized as basal textbook, linear skills, interest base, natural language, and integrated curriculum. During a two-year course of development, the instrument was revised to include common dimensions of teacher decision making that might be affected by various conceptual stances. These dimensions include criteria for judging pupils' success, criteria for forming instructional groups, allocation of time to reading activities, favored word-recognition prompts, emphasis on comprehension, and the teacher's view of the instructional role.

Examples:

> I believe that the teacher's role is to help children learn to love reading by allowing frequent free reading and by conducting individual book conferences.
>
> A B C D E
>
> I believe that all children should be systematically taught to use phonics skills.
>
> A B C D E
>
> I believe that reading groups should be based on pupils' interests.
>
> A B C D E

Validity, Reliability, and Normative Data: Work on The Proposition Inventory involved several stages of refinement. Initially the authors attempted to assess beliefs by writing propositions about reading on cards and asking teachers to sort the cards into piles of agreement and disagreement. Seventy items were field tested, after which the instrument was reduced to thirty-six items and administered again. Analyses and revision followed, resulting in the final forty-five-item form. The card sort format was rejected as inefficient, in favor of a series of Likert scales.

Six conceptual categories were represented during the early work: five reading theories (basal textbook, linear skills, interest base, natural language, and integrated curriculum) and a "confused/frustrated" category. The instrument was administered to graduate students at two universities. Factor analysis and reliability analysis were conducted. Factor analysis revealed that three, rather than the six intended, subscales were represented: (1) basal text and linear skills items; (2) interest-based, natural language, and integrated curriculum items; and (3) "confused/frustrated." Following these analyses, nondiscriminating items were revised or replaced and the "confused/frustrated" category was eliminated because of inability to validate it.

After the addition of items related to practical dimensions of decision making, the instrument was administered again. Factor analysis revealed two major subscales: a basal and linear skills orientation and an orientation toward natural language, interest-based and integrated models of instruction. A final revision involved rewriting certain individual items to improve their discrimination. The resulting form was administered to 128 students at Michigan State University. The reliability coefficients for the five intended subscales were computed, and a factor analysis revealed again that the interest, natural language, and integrated curriculum conceptions loaded on a common factor, while basal text and linear skills items loaded strongly on a separate factor. The final instrument was used to identify teachers for the Conceptions of Reading project, Institute for Research on Teaching, Michigan State University.

Ordering Information: ED 218 583

Related Document:

Duffy, G., and W. Metheny. *Measuring Teachers' Beliefs about Reading*. Research Series no. 41. East Lansing, Mich.: Michigan State University, Institute for Research on Teaching, 1979.

Category: Reading

Title: Comprehension Categories for Protocol Analysis

Author: William T. Fagan

Age Range: Primary to postsecondary

Description of Instrument:

Purpose: To interpret comprenension recalls in terms of textual information and cognitive processes used.

Date of Construction: 1981

Physical Description: Whereas a recall is a "product" of having read a text, it also results from two sets of processes: receptive and productive. The instrument outlines four stages for interpretive recalls. Stage one specifies which information is pertinent for analysis; stage two discusses the implications of which linguistic unit to use in categorizing the data. The comprehension categories are outlined in stage three, while stage four discusses the implications of weighting of the responses assigned to categories.

The procedures may be used with any recall. The first category and the criteria for assigning responses to this category are given below:

A. *Text Exact*
This category includes information from the text in its exact form or with minimal variations. It is assumed that this information is stored in rote fashion or is automatically constrained by other information and is "reproduced" in a similar state.

A1. *Verbatim Recall*
The information is a direct recall of the lexical items of the text.

Text: The boys were late for school.
Protocol: The boys were late for school.

Substitution of a determiner, a verb form, or a function word which does not change the meaning of the unit will also be placed in this category.

Text: He chased the animal.
Protocol: He chased an animal.

Text: People were waiting at the door.
Protocol: People were waiting by the door.

Text: The student had been absent many times.
Protocol: The student was absent many times.

A2. *Partial Recall*
A significant concept(s) (noun, verb, attribute) is/are omitted in the verbatim recall.

Text: After robbing the store, the convicts raced for their car.
Protocol: The convicts raced for their car.

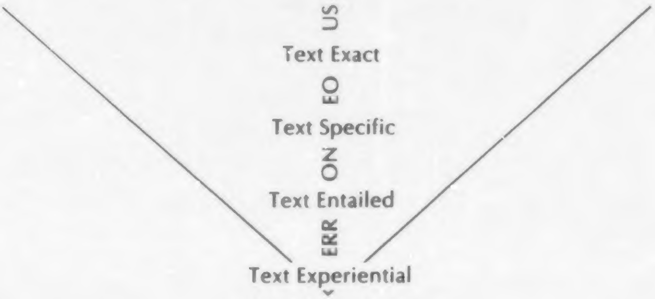
Text: The children had never seen such a tiny colt.
Protocol: The children had never seen such a colt.

This category would also include fragmented units that are not mazes and although not semantically complete, do indicate that the reader has noted and attempted to retrieve concepts which continue the story line.

Text: The stranger told him to follow his advice and put his lines at the spot indicated.
Protocol: The stranger told him . . . that he would put . . . all his lines . . .

Validity, Reliability, and Normative Data: The comprehension categories have been based on the construct of reading comprehension (as measured by a recall) as involving the reception and production of information that is generated from an interaction of the text data and the reader's prior knowledge. As indicated in the description of the categories, certain assumptions about the underlying processes that may contribute to that category are made based on the work of Kintsch and Van Dijk (1978).

The categories may be sequenced in terms of the proportion of text data and prior knowledge that may have contributed to the recall. This sequence may be illustrated by the following diagram with the amount of text decreasing from text exact to text experiential.



Interrater reliability based on assigning 187 clause units to categories produced the following percentages of agreement.

Category	A	98.5
	B	97
	C	93
	D	96
	E	95

Ordering Information: ED 236 664

Related Documents:

Fagan, William T. "Comprehension Categories for Protocol Analysis." Paper, University of Alberta, Edmonton, Alberta, Canada, 1981.

Kintsch, Walter, and Teun A. Van Dijk. "Toward a Model of Text Comprehension and Production." *Psychological Review* 85 (1978): 363-94. (EJ 191 792)

Category: Reading

Title: The Syntactic Proposition for Protocol Analysis

Author: William T. Fagan

Age Range: Primary to postsecondary

Description of Instrument:

Purpose: To divide recall protocols into units for subsequent category analysis.

Date of Construction: 1978

Physical Description: Various linguistic units have been used in the analysis of recall protocols. Such units include the sentence, T-unit, clause, phrase, and semantic proposition. The semantic proposition corresponds to a semantic idea and consists of a relation and one or more arguments. However, its application for dividing protocols is *abstract* in that the propositions in the protocol (protocol text base) are compared to the template text base that is the *idealized* representation of the meaning of a given text. An assumption underlying the syntactic proposition is that each idea unit that the subject intends to convey must be expressed syntactically. Conversely, it is assumed that by starting with the surface syntactic forms, these could be equated with semantic ideas.

A syntactic proposition may be of two types:

1. A *base syntactic proposition* is the simplest independent prediction (structurally) which may be used to convey information. Examples are:

(D) N V (Adj) (Adv)*(PP)*	Birds fly. The horse is black. Tom works hard. May smiles sweetly. The pony runs swiftly through the fields. Tom is diligent always in school.
(D) N V (Adv)* (N) (N)	Terry gave the dog a bone. The sergeant angrily gave the rookie a rebuke.

2. *Alternate syntactic propositions.* These are termed "alternate" to the base syntactic proposition because with the substitution or addition of words these could become base syntactic propositions.

*(Adv) may occur at various positions within the sentence.

*(PP) may not occur immediately after N.

The boy *with the books* is my brother.
There is a boy with books.
or
The boy has books.

He hid where the cook keeps the food.
The cook keeps the food there.

Twenty-three alternate syntactic propositions have been identified that may be subdivided into three groups: Relativization (Intact), Relativization (Ellipsis), and Complementation.

In the Relativization (Intact) group, all words are present in the structure although there may be a substitution (e.g., Have you seen the car *which* I bought?).

In the Relativization (Ellipsis) group, words must be added to alternate form to make up a base structure. These words are inferred (by the researcher) from the accompanying base structure.

Have you seen the car I bought?

The Complementation structures "complement" or "complete" a base structure.

The exercises are designed *to help you.*

Validity, Reliability, and Normative Data: Interrater reliability from several research projects using two independent raters has produced correlation coefficients ranging from .92 to .96.

Ordering Information: ED 236 636

Related Documents:

Brake, Delphine M. "An Investigation of Oral and Silent Reading with Low and High Achievers." Master's thesis. University of Alberta, Edmonton, Alberta, 1980.

Fagan, William T. *A Comparison of the Oral Language of Children, Ages Nine, Ten and Eleven.* Research Report. Canada Council Grant, S76-0563. University of Alberta, Edmonton, Alberta, Canada, 1978.

———. "A Longitudinal Study of Grade Five and Grade Eleven Students' Written Language and a Comparison to Teachers' Written Language." *Alberta Journal of Educational Research* 27 no. 3 (1981).

———. "A Longitudinal Study of Grade Five and Grade Eleven Students' Oral Language and a Comparison to Teachers' Oral Language." Paper. University of Alberta, Edmonton, Alberta, Canada, 1979.

Category: Reading

Title: The Quality, Direction, and Distance of within Sentence Contextual Constraints

Author: Victor Froese

Age Range: Grade 6

Description of Instrument:

Purpose: To investigate the types of responses given to a sentence completion task when constraint elements of word order, word form, redundancy, distance between lexical items, and interaction among these elements were considered.

This instrument differed from others in this area since it was designed to investigate the effects of lateral (forward and backward) constraints, it included the most potent forward constraining patterns from a previous study (Pike 1969) for cross-validation purposes, it controlled the sentence types used, and it included sentences of low and high contextual constraint generated by a systematic procedure.

Date of Construction: 1977

Physical Description: The instrument consisted of thirty-four sentences, half of which were high-associative sentences, and half of which were low-associative sentences. Each sentence was constructed so that the subjects were required to restore one word in each sentence (indicated by a line), although other words were omitted in order to study certain constraints. These words were designated by a line of x's (xxxxxxx). All sentences followed the same syntactic form: Determiner—Adjective 1—Noun 1—Verb—(Determiner 2)—Adjective 2—Noun 2. The constraints of one word form on another (and with different numbers of words intervening) were assessed. The base sentence (___: A) contained only one omitted word. (The ______ spider carried the dead fly.) In order to assess the constraining effect of Noun 2 on Adjective 1 (N2: A1), the sentence would look like this:

The ______ spider carried the dead xxxxx.

The seventeen constraints for one sentence pattern are listed below.

Sentence: A hot bath beats a cold shower.
01 —: A1 A _____ bath beats a cold shower.
02 N1: A1 A _____ xxxxxxxxxx beats a cold shower.
03 V: A1 A _____ bath xxxxxxxxxx a cold shower.
04 A2: A1 A _____ bath beats a xxxxxxxxxx shower.
05 N2: A1 A _____ bath beats a cold xxxxxxxxxx.

06 —:N1 A hot xxxxxxxxxx beats a cold shower.
07 V:N1 A hot _____ xxxxxxxxxx a cold shower.
08 A2:N1 A hot _____ beats a xxxxxxxxxx shower.
09 N2:N1 A hot _____ beats a cold xxxxxxxxxx.

10 —:V A hot bath _____ a cold shower.
11 A2:V A hot bath _____ a xxxxxxxxxx shower.
12 N2:V A hot bath _____ a cold xxxxxxxxxx.

13 —:A2 A hot bath beats a _____ shower.
14 N2:A2 A hot bath beats a _____ xxxxxxxxxx.

15 —:N2 A hot bath beats a cold _____.
16 A1:N2 A xxxxxxxxxx bath beats a cold _____.
17 N1:N2 A hot xxxxxxxxxx beats a cold _____.

Validity, Reliability, and Normative Data: The following procedures were used in constructing the instrument:

1. A subject-noun was arbitrarily selected from the stimulus words on Palermo and Jenkins's (1964) *Word Association Norms.* Examples would be "river," "cheese," and "doctor."
2. Next, the entry in *Word Association Norms* was scanned for the highest or lowest associates under the grade 4 to 6 headings. For example, "river" was selected for a LAS and hence the low-associative adjectives (two or fewer entries under the grade 4 to 6 columns) "muddy," "green," or "pretty," could be considered, or from the idiosyncratic responses, "turbulent," "moving," or "foam(ing)." Nouns of low-associative status could be "canoe," "falls," or "rapids," or, from the idiosyncratic responses in *Word Association Norms,* "rain," or "forest." The sentence resulting from the above possibilities was: "The foaming river carried the little canoe" (LAS).

   A similar procedure was followed in constructing the HAS sentences except only the highest associates under each entry were used. For example, "spider" was selected for a HAS and hence the high-associative adjectives possible were "big," "poison(ous)," and "dead." Possible nouns were "fly" and "web." The sentence resulting from these words was "The big spider carried the dead fly."

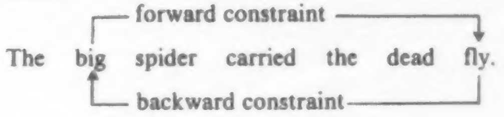
   Every effort was made to construct meaningful and sensible sentences within the repertoire of intermediate grade students. Some of the transitive verbs from the Palermo and Jenkins stimulus list were used, but more often the verb was arbitrarily selected to join the subject and object in a sensible manner.
3. The sentence paradigm used for this study was based on Riling's (1965) findings that the oral, written, and textbook language structures of sixth graders were similar, and that the subject-verb-object

pattern ranked first among them. For the present study, this high-frequency pattern was expanded by means of adjectives preceding the subject and object. The resulting pattern was as follows:

D	A1	N1	V	(D)	A2	N2
The	work	shoes	had		flat	heels.
The	pretty	woman	wore	a	lovely	dress.

The determiner in the second position was optional as indicated by the parentheses.

4. All adjective-noun pairs were checked by means of the *American Heritage Dictionary of the English Language* (1970) to avoid using compound nouns that would tend to inflate associativity. Several such pairs resulted during the preliminary stages of constructing the sentences and were subsequently omitted. Examples were bald eagle, wheel chair, machine gun, work bench, and rye bread.
5. Since forward and backward lateral constraints were of interest, two sets of deletion patterns were constructed. These two forms of contextual constraint may be conceptualized in the following manner:



No reliability data are given.

Ordering Information: ED 236 646

Related Documents:

Froese, Victor. "The Quality, Direction, and Distance of within Sentence Contextual Constraints." Ph.D. diss., University of Minnesota, 1977.

Palermo, D. S., and J. J. Jenkins. *Word Association Norms: Grade School through College*. Minneapolis: University of Minnesota Press, 1964.

Pike, L. W. "Structural Aspects of Contextual Constraints within Sentences." Ph.D. diss., Columbia University, 1969.

Riling, M. E. *Oral and Written Language of Children in Grades Four and Six Compared with the Language of Their Textbooks*. Cooperative Research Project No. 2410. Durant, Okla.: Southeastern Oklahoma State College, 1965. (ED 003 022)

Category: Reading

Title: Feedback to Oral Reading Miscue Analysis System

Authors: James V. Hoffman
Christopher J. Baker

Age Range: Primary to postsecondary

Description of Instrument:

Purpose: To classify miscue-focused verbal interactions between teacher and pupil(s) during oral reading instruction.

Date of Construction: 1980

Physical Description: The FORMAS taxonomy includes five major clusters of teacher/pupil interactive behaviors. The first cluster deals with the actual miscue and provides for an analysis of the miscue type (omission, substitution, etc.) and its characteristics (graphophonic similarity, etc.). Cluster II, Student Reaction to Miscue, focuses upon student behaviors that are manifested subsequent to a reading miscue and prior to teacher intervention. Teacher Verbal Feedback constitutes Cluster III. Four types of feedback are identified within this cluster. The fourth cluster, Student Verbal Feedback, records the involvement of other students present during the interaction while the final cluster, Miscue Outcome, focuses on what happens ultimately to the observed miscues. This cluster specifies who finally identifies a miscue.

The starting point for the FORMAS is a tape-recorded sample of oral reading instruction. This sample can be a group lesson or one in which a teacher and single student are working together. The teacher then locates each miscue on the tape and its text stimulus. The taxonomy can be represented on a chart or coding sheet. Miscues are written in at the far left and the five-cluster analysis of each miscue is then checked off, moving from left to right across the coding sheet. An example of a coding sheet inclusive of Cluster I is presented on the following page.

Validity, Reliability, and Normative Data: Interrater reliability coefficients for agreed upon miscues among trained coders has typically ranged between 0.83 and 1.00.

Ordering Information: See related documents.

Related Documents:

Hoffman, James V., and Christopher J. Baker. "Characterizing Teacher Feedback to Student Miscues During Oral Reading Instruction." *The Reading Teacher* 34 (1981): 907-13. (EJ 245 507)

Hoffman, J. V., C. Gardner, and R. Clements. *Training Manual for FORMAS—Feedback to Oral Reading Miscue Analysis System.* Technical Report #5085. Austin: Research and Development Center for Teacher Education, University of Texas. (ED 192 291)

			Cluster I Miscue										
			Type							Characteristics			
Miscue No. (order of occurrence)	Expected Response (and its location in the text)	Observed Response and ID no. of student reading)	Insert	Omission	Substitution	Mispronunciation	Don't know	Hesitation	Repetition	Little change in meaning	Substantial change in meaning	High graphophonic similarity	Low graphophonic similarity
1	rescued (5)	saved (1)			X					X			X
2	A (20)	one (1)			X					X			X
3	the (40)	A (1)			X					X			X
4	lonely (50)	alone (1)			X					X			X
5	exhaust (69)	— (1)						X					
6	wasn't (76)	was (1)		X							X		
7	penetrate (85)	pentrate (1)				X					X	X	

Category: Reading

Title: Textbook Usage Inventory (TUI)

Author: Indiana State Department of Public Instruction

Age Range: Grades 4 through 12

Description of Instrument:

Purpose: The Textbook Usage Inventory is designed to assess students' ability to read required materials used in a given content area course and to serve as a helpful tool for the content area teacher in the implementation of an approach to teaching that will best communicate the content to all students.

Date of Construction: June 1981

Physical Description: Guidelines are provided for constructing a TUI that can be easily adapted for any content area. Initially a checklist of reading skills and resources is given from which teachers may choose those items which are relevant to their particular content. Skills and resources are listed for the following categories: literal, interpretative, critical, creative, classroom resources, and library resources. Teachers are then provided with a basic outline for constructing a TUI and examples of inventory items.

Eight specific steps for administering the TUI are indicated. In general, teachers should explain the purpose of the inventory to the students: that it represents an effort on the part of the teacher to become aware of individual needs and will not constitute part of their grade. Two evaluative criteria are obtained: (1) percentages indicate whether a student is functioning effectively with a textbook. If the material is suitable, scores should range between 70 and 90 percent. (2) By coding the types of questions asked, a student's comprehension strengths and weaknesses may be noted. A student is considered to be deficient in any one specific skill if he or she answers more than one question incorrectly when there are at least three questions measuring a specific skill.

Validity, Reliability, and Normative Data: Content validity may be claimed for the inventory since the items chosen for evaluation are based on the specific course taught.

Ordering Information: ED 210 648

Category: Reading

Title: Extended-Cloze Reading Skills Test of Hierarchical and Spatial/Chronological Ordering Ability

Author: Anne V. Martin

Age Range: Adults (ages eighteen through thirty-five) and possibly high school

Description of Instrument:

Purpose: To determine readers' ability in processing information relationships in written discourse in English. It is particularly appropriate for ESL students but may also be used with dialect speakers.

Physical Description: The extended-cloze instrument consists of six reading passages (230-300 words long), each with four sentences deleted from the written text (a total of twenty-four items). Extended-cloze is defined as deletion of entire sentences, rather than of individual words as in traditional cloze procedures, from unified discourse in written English. Students read each passage and select the best of four multiple-choice alternative sentences for each deletion. The instrument contains two subtests: three of the passages measure processing of hierarchical order; three measure processing of spatial/chronological order. Hierarchical (H) order is defined as a developmental sequence of concrete or abstract ideas in written English, consisting of steps along a continuum of general to specific information. Spatial/chronological (S/C) order is defined as a developmental sequence in written English in which objects and/or events are related in space and/or time.

The sample item with instructions is given below:

> This exercise consists of six fictional reading passages with deleted sentences. Deletions are indicated with a long line. There are four possible sentences given for each deletion.
> Each time, read the entire passage. Then go back to the first deletion. Read the four possible answers. Choose the *best* one and circle its letter *on the answer sheet* that is provided. Go on to the next deletion in the passage. After you finish the items in the passage, go on to the next passage.
>
> "Fairfield has many green areas within its city limits. One of the most beautiful places is Wilson Park. ____. People from all over Fairfield visit the park and enjoy themselves."
>
> a. The North Fork River is another spot.
> b. There are other green spots in and near the city of Fairfield.
> c. It has picnic grounds and wide open grassy spaces.
> d. The statue near the parking lot was created by a French artist.

Validity, Reliability, and Normative Data: Construction of the passages was based on five design criteria:

1. All contextual locations are fictional and as nonculture-specific as possible. Concepts used in the passages (e.g., satellite, boat, college campus) are familiar to students for whom the instrument is suitable, both native and nonnative speakers of English.
2. Vocabulary items do not pose barriers to comprehension. In addition to extensive screening using ESL students, the vocabulary items were checked against the Brown Corpus (Kučera and Francis 1967). All but 2.7 percent (nineteen words) are high-frequency terms, and in context the remaining nineteen words (e.g., *canoe, sofa*) do not pose comprehension problems for students.
3. Readability levels were measured using the Fry Readability Scale (Fry 1968). All passages range from mid-tenth- to low-twelfth-grade level, with three of the passages at the eleventh-grade level.
4. Sentence structure complexity was controlled to minimize the number of embeddings, relative clauses, verb forms, and verb tenses, but to retain natural-sounding language. Mean length of sentences is sixteen words (range six to twenty-seven words).
5. Obvious transitional expressions such as "First" and "The second type is . . ." that might provide clues to the appropriate response were omitted from items and passages, except where essential to the flow and coherence of written English.

In constructing the response choices for the H items, all four alternatives to each item were selected so that they were contextually appropriate and "equally plausible"; they differ in hierarchical appropriateness. Each deleted sentence (keyed response) is a step on the continuum of general-to-specific information in the passage. For each item there are three distractors:

1. too general for that point in the context,
2. too specific for that point in the context,
3. irrelevant, a violation of native-speaker expectation of appropriate information for that position, taking into account information both preceding and following the deletion.

For two items, discourse constraints made one distractor of type one and one of type two impossible; in these cases there are two of either type one or two.

For each of the S/C items, all four alternatives for each item are appropriate to the context, but the three distractors are inappropriate spatially or chronologically within the particular discourse setting. That is, a native speaker of English would expect reference to a certain spatial or temporal event (as given in the keyed answer) to follow information preceding the deletion and to precede information in subsequent sentences.

An item analysis based on the responses of 438 college students (212 native speakers and 226 nonnative speakers of English) showed point biserial coefficients for all items within each subtest [r_{pbis} (436) $\geq$ .196] with $p > .05$. The phi coefficients range (.19 − .44) all exceeded .01 level using chi-square as the test of significance.

A factor analysis confirmed cross-cultural validity (the same underlying factor structure for native and nonnative speakers of English), and the presence of two subtests associated with items designated as H or S/C. Through factor analysis also, construct validity was established for spatial/chronological ability. A factor interpretable as hierarchical ordering ability was not identified, although it was established that the H section measures some distinct underlying construct(s).

Reliability based on the Kuder-Richardson formula pooled for the 438 subjects yielded coefficients of .76, .55, and .70 respectively for the twenty-four-item instrument, the twelve-item H subtest, and the twelve-item S/C subtest.

Ordering Information: ED 236 640

Related Documents:

Fry, E. "A Readability Formula that Saves Time." *Journal of Reading* 11 (1968): 513, 575-78.

Kučera, H., and W. N. Francis. *Computational Analysis of Present-Day American English.* Providence, R.I.: Brown University Press, 1967.

Martin, A. V. "Proficiency of University-Level Advanced ESL Students and Native Speakers of English in Processing Hierarchical Information in Context." Ph.D. diss., University of Southern California, 1980. (Available from Micrographics Department, U.S.C., Los Angeles, CA 90007.)

Category: Reading

Title: Morpheme Facilitation Test

Author: David W. Moore

Age Range: Intermediate

Description of Instrument:

Purpose: The Morpheme Facilitation Test estimates the effect of the morphemic composition of words on students' identification of those words.

Date of Construction: 1980

Physical Description: The Morpheme Facilitation Test includes thirty-six sentences that each contain one target word. The target words vary according to three levels of morphemic composition: (1) monomorphemic words (e.g., *armadillo*), (2) suffixed words with a spelling change in the stem (e.g., *competition*), and (3) suffixed words with no spelling change in the stem (e.g., *betrayal*). All stems are free forms that can stand alone as English words. The sentences are typed on separate sheets of paper which are shuffled and then handed to a student. The student reads each sentence orally in turn and hands back the sheet. Whether each target word is pronounced accurately or not is recorded on an examiner's protocol. No time limit is set. The raw score of correct responses for words of each of the three types of morphemic composition is then totalled and compared. Testing requires 10–15 minutes for each student. The complete test is given below.

Morpheme Facilitation Test Sentences

Monomorphemic Words

A pet *armadillo* is quite unusual.
The sudden *avalanche* destroyed the town.
The old *derelict* begged for money.
Jim's new *enterprise* earns much money.
Next week's *episode* should be exciting.
My brother's *fantasy* made me laugh.
Last year's *hurricane* was the worst.
Our group *intellect* will solve this.
The wide *panorama* left Sharon speechless.
This next *semester* will be easy.
The terrible *tragedy* was on T.V.
The rocket's *velocity* worried the men.

Suffixed Words—Spelling Change in the Stem

Next week's *competition* will be easy.
This toy's *creator* lives next door.
All her *devotion* is for Jimmy.

Their new *happiness* should last forever.
The worst *hindrance* is the traffic.
The enemy's *hostility* was easily seen.
His strong *immunity* keeps him healthy.
My mother's *permission* was a surprise.
The band's *pianist* plays quite well.
Bill's new *promotion* really pleased him.
A cold *sensation* crept through him.
His parking *violation* cost ten dollars.

Suffixed Words—No Spelling Change in the Stem

His friend's *betrayal* deeply hurt him.
One more *deduction* will ruin me.
Their next *engagement* is in Atlanta.
The marching *formation* was single file.
The captain's *heroism* inspired his men.
Our English *instructor* likes to talk
His coaching *mastery* won many games.
That one *prisoner* looks really mean.
This country's *prosperity* is in danger.
The next *settlement* is named Parker.
The young *socialist* lives in Russia.
A plant's *tendency* is toward sunlight.

Validity, Reliability, and Normative Data: The content validity of this instrument is revealed by the morphemic composition of the target words. Twenty-four words contain more than one morpheme and twelve contain only one. Of the twenty-four words that contain more than one morpheme, twelve include stems with spelling changes and twelve include stems with no spelling changes. Care was taken to select words that varied according to levels of morphemic composition but that were controlled for various properties that contribute to word identification. The three sets of twelve words each were made equal according to six properties: (1) imagery rating, (2) frequency, (3) word length, (4) number of syllables, (5) pronounceability, and (6) form class. In addition, a uniform, six-word sentence context was employed so that the target word always appeared as the sentence subject preceded by two modifiers. The thirty-six sentences were tested to insure that they provided sufficient contextual information for each target word so that readers consistently produced syntactically acceptable responses. However, the sentences were also tested to insure that advanced readers could not consistently produce the exact target word on the sole basis of the contextual information.

The reliability of this test instrument was determined by two procedures. In the first procedure, the internal consistency of students' responses at each morphemic level was assessed by the Kuder-Richardson 21 statistic. (Coefficients ranged from .77 to .84.) The second procedure correlated test scores across levels of morphemic composition with Pearson correlation coefficients that ranged from .84 to .85 and were statistically significant.

Data from a sample of third- through sixth-grade students showed that they identified slightly more suffixed words ($\bar{x} = 8.15$) than monomorphemic words ($\bar{x} = 7.18$). Practically no differences were obtained between suffixed words with a spelling change ($\bar{x} = 7.98$) and suffixed words with no spelling change ($\bar{x} = 8.23$).

Ordering Information: Full instrument reproduced above.

Related Document:

Moore, D. W. "The Effect of Morphemic Composition of Words and Grade Level of Students on Word Identification." Ph.D. diss., University of Georgia, 1980.

Category: Reading

Title: An Assessment of the Effects of Different Error Types on the Understanding of Connected Discourse

Author: Tom Nicholson

Age Range: Primary, upper elementary

Description of Instrument:

Purpose: To analyze systematically the relative effects of different types of oral reading errors on comprehension.

Date of Construction: 1977

Physical Description: This instrument was considered unique since it allowed for investigating the effects on comprehension of five factors frequently reviewed in the literature: error type (here referred to as "simulate type"), set strength, passage difficulty, error rate, and access to text. The instrument consisted of a basic set (each with an easy and a difficult version) of six stories. Every story was transformed so that it contained simulated errors of a particular type: correct (CORR), semantically related visually unrelated (SRVU), semantically unrelated visually related (SUVR), semantically and visually unrelated (SUVU), no response (NONE), and the mixed simulate type version (MIX). There was a high-rate condition in which fifteen of the nouns were replaced by different simulate types, whereas six were replaced in the low-rate condition. In each story, set strength was established by repeating all nominals with comprehension probes, in their correct form, at least once. The high set strength condition was established by first repeating the target word in its correct form later in the passage, and then using the correct form a second time as part of a title for the story. Thus, the reader was given two appropriate cues to the real meaning of the initial simulate. In the low set strength condition, the nominals being tested were repeated only once in the passage in their correct form. After reading each story, the subjects answered six cloze-type questions (with deleted target constituents) and two additional questions designed to test inferential comprehension. Examples are given below of an original story, easy in difficulty, in high set strength condition, and in semantically related visually unrelated (SRVU) and semantically unrelated visually related (SUVR) versions.

> *The newspaper, the home, and the money*
>
> One day Russell and his friend Timothy were playing after school. Finally they decided to go home. They walked along the edge of a high wall. Then they jumped over a fence. Then they started walking backwards. Because he couldn't see where he

> was going, Russell bumped into Mr. Zinder's newsstand. He knocked his money box off the counter. The money rolled all over the road. Mr. Zinder started yelling. He was waving a newspaper wildly in the air. The boys ran off, jumped a nearby fence, and went home. Now both boys avoid buying a newspaper from Mr. Zinder's newsstand.
>
> *The newspaper, the home, and the money*
>
> One time, Russell and his playmate Timothy, were playing after class. Finally they decided to go cabin. They walked along the side of a high ledge. Then they jumped over a barrier. Then they started walking backwards. Because he couldn't see where he was going, Russell bumped into Mr. Zinder's table. He knocked his money case off the table. The coins rolled all over the street. Mr. Zinder started yelling. He was waving a magazine wildly in the sky. The students ran off, jumped a nearby fence, and went home. Now both boys avoid having to buy a newspaper from Mr. Zinder's newsstand.
>
> *The newspaper, the home, and the money*
>
> One dam, Russell and his frog Timothy were playing after port. Finally they decided to go horse. They walked along the egg of a high wave. Then they jumped over a face. Then they started walking backwards. Because he couldn't see where he was going, Russell bumped into Mr. Zinder's nonsense. He knocked his money boot off the cradle. The mountains rolled all over the rope. Mr. Zinder started yelling. He was waving a neighbour wildly in the ant. The books ran off, jumped a nearby fence and went home. Now both boys avoid having to buy a newspaper from Mr. Zinder's newsstand.

Validity, Reliability, and Normative Data: In order to assess the influence of various factors on comprehension, it was decided to experimentally simulate these factors in reading but in a way that is amenable to rigorous statistical analysis. In this design, a number of stories were transformed so as to simulate the reading environment in which the unskilled reader finds himself or herself when answering comprehension questions. The transformed stories were then given to proficient readers who tried to read and understand the anomalous material. In a sense, the proficient readers were required to complete certain comprehension tasks in an environment similar to that faced by beginning readers.

Certain steps guided the selection of the stories. Each story had to be 100 words in length, with only fifteen nominals for each story. Only narrative-type stories, involving a chain of causal events leading to some kind of resolution, were included. Each story was classified as either easy or hard. Then each story went through each rate of simulate substitution: the high-rate version (15) and the low-rate version (6). The story was then transformed to correspond to each of the simulate types. The classification scheme devised for generating the artificial error types (simulates) was based on the interaction of two factors assumed to be independent: visual features and semantic characteristics of words. The interaction can be represented by a two-item by two-item table, where words can be classified according to their degree of visual and semantic relatedness to the target word.

		Semantic Relation	
		Related	Unrelated
Visual Relation	Related	house	horse
	Unrelated	cottage	apple

Three general constraints and six specific constraints were specified for the selection of each simulate type (Nicholson 1977, 60–62). In addition, four constraints were specified (66) for formulating response choices for the inferential questions. Finally, ten levels of semantic appropriateness were defined for classifying responses to the cloze items (81–84), and a five-level scale was devised for scoring responses to inference questions (85).

It was assumed that since the categories were highly specific and strictly defined, reliability of classification of responses would be high. In order to strengthen the reliability of the classification system, the analyses were carried out on two levels from more broadly based to specific measures.

Ordering Information: ED 236 637

Related Document:

Nicholson, Tom. "The Relative Effects of Different Error Types on Understanding of Connected Discourse." Ph.D. diss., University of Minnesota, 1977.

Category: Reading

Title: A Tentative Criterion-Referenced Test to Measure Thinking Processes, Form A and B

Author: Jo Ellen Oliver

Age Range: Ages seven to eleven

Description of Instrument:

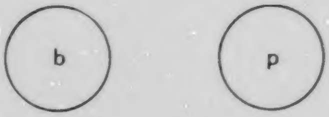
Purpose: The test was devised to measure children's abilities to synthesize concepts from several sources.

Date of Construction: 1978

Physical Description: An assumption was made that children move from the concrete to the abstract, from specific to generic, and that the ability to synthesize and form new concepts is an important prerequisite to reading comprehension. Each form of the test is divided into three parts with part one focusing on letters, part two on words, and part three on stories. Children's responses to these items are labeled as perceptual or conceptual, with conceptual scores indicating a more developed ability to generalize and synthesize. The test is written on a primer level of difficulty, is easy to score (scoring guidelines are included), and requires about 30 minutes to complete. The test may be administered individually or in groups.

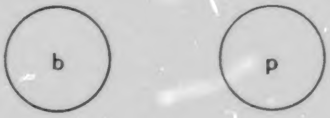
Sample items from each of the three sections of Form A are given below:

I. A. Look at what is in the circles.
How are they alike?



1.
2.

B. Look at what is in the circles.
How are they different?



1.
2.

II. A. Look at what is in the circles.
How are they alike?

come came

1.
2.

B. Look at what is in the circles.
How are they different?

come came

1.
2.

III. (With reference to two stories, "The Ant and the Grasshopper" and "The Little Red Hen," the following questions are asked):

A. 1. Who worked hard in the story "The Little Red Hen"? Who worked hard in the story "The Ant and the Grasshopper"?

2. Were there some animals in both of the stories who were lazy?

3. Who got to eat in both of the stories?

B. How are these two stories alike?

1.

2.

C. How are these two stories different?

1.

2.

D. 1. How do you feel about the way the hen and the ant acted?

2. How do you feel about the way the other animals acted?

3. Describe which animal in the stories you would like to be and tell how you would have acted.

4. What did you learn from the stories?

Validity, Reliability, and Normative Data: The form of the test (criterion-referenced) was inspired by Otto (1973), who in his discussion of criterion-referenced and norm-referenced tests, said of criterion-referenced tests, "they get directly at the performance of individuals with regard to specified individual objectives" (p. 18).

The first draft of the test was based on criteria established by Stauffer (1969) for consideration in the construction of an instrument to measure thinking processes.

The test was tried on a number of children and then submitted for scrutiny to two college professors and a seminar of thirty doctoral students in reading. Their suggestions were incorporated into a revised form and its alternate form. These forms were then submitted for evaluation to four college professors, with backgrounds in testing and child development, to establish content validity. To establish partial construct validity, the test or its alternate was administered randomly to thirty fifth-grade children. These children were then given 90 minutes of training in the conceptualization process. After eight days had elapsed, the children were given the alternate form of the test they had taken previously. A paired *t* test was performed. For the total scores, the *t*-value was 2.41 ($p < .05$). For conceptual scores only, the *t*-value was 3.23 ($p < .005$). Mean total points increased 2.50 and mean conceptual points increased 2.6.

To establish cross-form reliability, 100 fifth graders in a different school were randomly given Form A or Form B. Twelve days later, the children were given the alternate forms of the tests they had taken. Children were given no training. Cross-form reliability (based on a Pearson product-moment correlation) for the total test score was .80 ($p < .001$). For the conceptual scores on the tests, cross-form reliability was .74 ($p < .001$). The correlations between the two scores of the tests, conceptual and total, were .89 ($p < .001$).

Interrater reliability based on the independent scoring by two raters was .99 for the total test scores and .99 for the conceptual scores.

Ordering Information: ED 236 645

Related Documents:

Oliver, Jo Ellen. "The Effects of Training in Synthesis-Level Conceptualization upon Low-Achieving First and Fifth Graders Involved in a Literature-Based Cross-Age Tutoring Program." Ph.D. diss., University of Georgia, Athens, 1978.

Otto, W. "Evaluating Instruments for Assessing Growth in Reading." In *Assessment Problems in Reading,* edited by W. H. MacGinitie. Newark, Del.: International Reading Association, 1973.

Stauffer, R. G. *Teaching Reading as a Thinking Process.* New York: Harper & Row, 1968.

Category: Reading

Title: The Standard Test of Reading Effectiveness (STORE)

Author: Elray L. Pedersen

Age Range: Upper elementary, secondary, college, and continuing education students.

Description of Instrument:

Purpose: To assess a reader's achievement in various reading skills for instructional placement.

Date of Construction: 1980

Physical Description: The test contains three forms (A, B, C), each of which contains fifty items. The items test comprehension of idioms and direct statements, but place emphasis on drawing inferences. Administration time is 25 minutes. Two sample items are given below.

> I fell and broke my arm. I told the doctor I was experiencing:
>
> A—pain B—joy C—surgery D—ill health
>
> A wise old owl sat in the branches of an oak,
> The longer he sat, the less he spoke
>
> A. The less he spoke, the more he learned.
> B. The longer he sat, the longer he sat.
> C. The less he spoke, the less he spoke.
> D. The less he spoke, the more he spoke.

Validity, Reliability, and Normative Data: None are available to date.

Ordering Information: ED 236 669

Category: Reading

Title: Categories of Inferencing Strategies

Author: Linda Phillips-Riggs

Age Range: Primary to postsecondary

Description of Instrument:

Purpose: To describe which strategies subjects use while making inferences.

Date of Construction: 1981

Physical Description: Ten inferencing strategies have been identified. These are:

1. Rebinding.
2. Questioning a default interpretation and/or a direct or indirect conflict.
3. Shifting of focus.
4. Analyzing alternatives.
5. Assigning an alternate case.
6. Confirming an immediate prior interpretation.
7. Confirming a nonimmediate prior interpretation.
8. Assuming a default interpretation and transforming information.
9. Neglecting to respond or holding information.
10. Empathizing from experience.

A definition and example for strategy one—rebinding—is given below:

> Definition: When a reader suggests or hypothesizes a possible interpretation and then *immediately* realizes that this interpretation conflicts with previous information, he/she then substitutes another interpretation. In essence, the reader is rebinding the present information to a previous interpretation.
>
> Example:
> *Passage on Farming*
> "They're using the auger to put the wheat into the graineries—no, not the graineries, but their bins."
>
> *Passage on Going to the Rodeo*
> In response to the question "Why would Marty need glasses?" "Marty has glasses and he forgot them at home—no, he lost them."

A reader is asked to read a passage up to a particular point at which an inference is to be made (predetermined by the examiner). The subject is then asked to tell what is happening and what may happen next. If necessary, probing questions are asked. The responses are then analyzed in terms of the strategies that are being used.

Validity, Reliability, and Normative Data: A theory of reading, within which one of the components was *encounter* (between reader and text), was established as a construct in which to specify the strategies used (that is, the nature of the encounter). Identification of the strategies was based on research by Collins et al. (1977). Two pilot studies were conducted to streamline procedures and to isolate the strategies used. Agreement was necessary between the researcher and a colleague. Data from forty sixth-grade readers were then analyzed. A factor analysis indicated that all correlations were less than .20 and that the pattern of factor loadings clearly identified a single factor for each strategy. Each strategy was obviously independent of the others.

Interrater reliability between the researcher and a colleague on one-tenth of the protocols showed agreement of 93.4 percent.

Data on the use of the strategies by forty sixth-grade children and the influence of specific variables are given in the related document by the test's author.

Ordering Information: ED 236 667

Related Documents:

Collins, A., J. S. Brown, and K. M. Larkin. *Inference in Text Understanding.* Technical Report No. 40. Urbana: University of Illinois, Center for the Study of Reading, 1977. ED 15υ 547

Phillips-Riggs, Linda. "The Relationship between Reading Proficiency, Background Knowledge, and Inferencing Strategies." Ph.D. diss., University of Alberta, Edmonton, Alberta, Canada, 1981.

Category: Reading

Title: Comprehension Process Score

Author: Mark C. Sadoski

Age Range: Preschool through adult

Description of Instrument:

Purpose: To categorize a reader's miscues according to their predictability of levels of comprehension.

Date of Construction: 1980

Physical Description: The procedure may be used with any oral reading data. The oral reading is taped and scripted according to the guidelines of the *Reading Miscue Inventory Manual* (Y. Goodman and Burke 1972). The semantic acceptability and correction behavior of each of twenty-five consecutive miscues is recorded. Percentages are calculated for supercues, pseudocues, and entropicues. Supercues are defined as: (1) miscues that are semantically acceptable in the entire passage and are successfully corrected; (2) miscues that are semantically acceptable in the entire passage and for which no correction is attempted; and (3) miscues that are semantically acceptable with the portion of the sentence up to the point of the miscue and are successfully corrected. Pseudocues are miscues that are semantically unacceptable and for which no correction is attempted. Entropicues are all other miscues not classified as supercues or pseudocues (Carey 1978). To score, the percentage of entropicues is subtracted from 100 percent, leaving a residual percentage. The percentage of supercues is divided by the residual percentage to yield a ratio. This ratio is multiplied by 100 to yield the comprehension process score (Sadoski 1970). Scoring range: 0–100. The complete instrument is given in Figure 1.

Validity, Reliability, and Normative Data: The comprehension process score is theoretically based on the concept of the altercue continuum (Page 1978) that categorizes miscues as either supercues (predictors of high comprehension levels), pseudocues (predictors of low comprehension levels), or entropicues (nonpredictors of high or low comprehension levels). Empirical evidence for the altercue continuum was established by Carey (1978), yielding operational definitions of the three altercue types.

The construct validity of inferring reading comprehension from the semantic acceptability and correction behavior exhibited on miscues is established by the fact that these responses indicate when readers are or are not reconstructing the meaning of an author's message (K. Goodman 1969). The concurrent validity of the comprehension process score was

Subject: ________
Story: ___________
CP Score

Supercues (total % of 004,001,009) = 100% (% of supercues)
Pseudocues (total % of 020) = −____ (% entropicues) ____ × 100 =
Entropicues (total % of remaining) = = residual (residual)

Correction Behavior

Semantic Acceptability	successful	partially successful	not successful	no attempt
passage	001	002	003	004
sentence	005	006	007	008
prior context	009	010	011	012
after context	013	014	015	016
not acceptable	017	018	019	020

Figure 1. Comprehension Process Score Instrument (Sadoski 1980).

described by Sadoski, Page, and Carey (1980), who demonstrated substantial and significant correlations to post-oral reading cloze test scores and to standardized reading comprehension test scores.

The reliability of the comprehension process score was demonstrated by Sadoski (1981): randomly selected comprehension process scores were recalculated and found to correlate with original scores at the .96 level.

Ordering Information: Full instrument reproduced above.

Related Documents:

Carey, R. "A Psycholinguistic Analysis of the Effects of the Semantic Acceptability of Oral Reading Miscues on Reading Comprehension." Ph.D. diss., University of Connecticut, 1978.

Goodman, K. "Analysis of Oral Reading Miscues: Applied Psycholinguistics." *Reading Research Quarterly* 5 (1969): 9–30.

Goodman, Y., and C. Burke. *Reading Miscue Inventory*. New York: Macmillan, 1972.

Page, W. "The Altercue Continuum: Theoretical Considerations." Paper presented at the Buffalo Conference on Researching Response to Literature and the Teaching of Literature, Buffalo, N.Y., 1977. (ED 155 677)

Sadoski, M. "A Unitive Psycholinguistic Comprehension Measure: The Comprehension Process Score." In *Language Centered Reading Instruction*, edited by J. Meagher and W. Page. Storrs, Conn.: University of Connecticut, Reading-Language Arts Center, 1980.

———. "Reliability Estimates for Comprehension Process Scores." Paper, University of Connecticut, Reading-Language Arts Center, Storrs, Conn., 1981.

Sadoski, M., W. Page, and R. Carey. "The Comprehension Process Score: Empirical Testing of a New Miscue Scoring Technique." Paper, University of Connecticut, Reading-Language Arts Center, Storrs, Conn., 1980.

Category: Reading

Title: Smith/Palmer Figurative Language Interpretation Test

Authors: Edwin H. Smith
Barbara C. Palmer

Age Range: Intermediate to postsecondary (including adult basic education students)

Description of Instrument:

Purpose: To assess the ability to interpret the major types of figurative language or tropes such as similes, metaphors, proverbs, and personification.

Date of Construction: 1979

Physical Description: The figures of speech used in the Smith/Palmer Figurative Language Interpretation Test items were selected from those found in basal reader series, Sutherland's (1973) recommended trade books, Brewer's *Dictionary of Phrase and Fable* (1978), and Bartlett's *Familiar Quotations* (1968), plus newspapers and magazines such as *Time* and *Newsweek*. The pilot version of the test contained two equal parts of fifty items each—part one tested the meaning of the figures of speech in isolation; part two tested their meanings in the context of a sentence(s). (See the sample items at the end of this section.) In both parts, the reader was required to select which of four possible answer choices represented the most common meaning for each figure of speech. The test was not timed. To minimize other comprehension difficulties, all test items were written at or below the fifth-grade readability level, using the EDL Revised Core Vocabulary (1969). All figurative statements were also excluded from answer choices. Also, the figurative statements were underlined as an aid to the reader. At the end of the pilot study, based on a detailed item analysis, the instrument was revised and the 100 improved items were randomly divided into Form A and Form B with 50 items each. Within each of the tests, the original part one/two, multiple-choice format was retained.

Part I:

Walking on air means feeling:
A. impatient
B. careful
C. joyful
D. careless

Part II:

Just by listening to the two of them, you could see that they were painted with the same brush. They were:

A. very much alike
B. both young
C. sisters
D. relatives

Validity, Reliability, and Normative Data: Content validity was established for the Smith/Palmer Figurative Language Interpretation Test by selecting a large number of figures of speech from basal reader series and trade books; these were supplemented by selections from the works of Sutherland, Brewer, and Bartlett. These sources were chosen because of their common usage. In addition, the content and format were compared with similar instruments such as those found in the dissertations of Muller (1976) and Hartman (1978). After a detailed item analysis, the 100-item test was revised and the improved items were randomly divided into two forms of the test, with fifty items each.

Reliability was assessed by the Kuder-Richardson Formula 20, which on three occasions yielded coefficients of .83, .84, and .76.

The mean performance on the 100-item test for a sample of grade 7, 8, and 9 students was 40.40 with a standard deviation of 10.61 and a standard error of measurement of 4.43. For a group of teachers enrolled in a graduate reading extension course, the mean was 82.45 with a standard deviation of 4.55 and a standard error of 3.11. The administration of each form (fifty items) with adult basic education students yielded a mean of 27.89 with a standard deviation of 7.59 and a standard error of 3.06 for Form A, while for Form B the statistics were 27.80, 6.11, and 3.01.

Ordering Information: ED 236 668

Related References:

Bartlett, John. *Familiar Quotations.* 14th ed. Boston: Little, Brown, 1968.

Brewer, E. Cobham. *The Dictionary of Phrase and Fable.* New York: Avenel Books, 1978.

Hartman, Sharon D. "An Investigation of the Precision of Metaphorical Language Interpretation of Students in Grades Six and Nine." Ph.D. diss., Florida State University, 1978.

Muller, Dorothy H. "An Investigation of the Precision of Metaphorical Language Interpretation of Students in Grades Four through Seven." Ph.D. diss., Florida State University, 1976.

Smith, Edwin H., and Barbara C. Palmer. *Figuring Out Figurative Language.* Project AS9-13: Implementing a Model for Increasing Reading Comprehension of Adult Basic Education Students, Tallahassee: Florida Department of Education, 1979.

Sutherland, Zena, ed. *The Best in Children's Books: The University of Chicago Guide to Children's Literature 1966–1972.* Chicago: University of Chicago Press, 1973. (ED 085 767)

Category: Reading

Title: Degrees of Reading Power (DRP)

Author: The State Education Department, New York.

Age Range: Upper elementary

Description of Instrument:

Purpose: (1) To measure the most difficult text that a student can read with comprehension; (2) to evaluate the outcome of reading instruction or to gauge an individual's growth over a relatively short period of time; (3) to match instructional materials to the student's reading capabilities.

Physical Description: The DRP consists of a series of nonfiction prose passages on topics randomly selected from the *Encyclopaedia Britannica.* The passages are presented in order of difficulty beginning with very easy text and progressing by degrees to very difficult text. Test items are created by the deletion of seven words in each passage, with the student selecting the most appropriate word from the five options provided for each deletion. In order for the student to answer DRP questions correctly, he or she must read and comprehend the text pertaining to those items. This requirement represents a clear contrast with *some* conventional reading tests in which questions can sometimes be answered without reading the pertinent text. The DRP has been scaled with the Rasch model, and the "average" item difficulty per passage has been correlated with the Bormuth readability formula. Therefore, it is possible to relate a student's total number correct on the DRP to a set of predictions concerning the most difficult prose reading materials that the student can comprehend with given probabilities. The probability levels that can be reported are: the independent level, where $p = .90$; the instructional level, where $p = .75$; and the frustration level, where $p = .50$. This information will enable the teacher to assign materials to students based on their probability of success at handling the materials.

The following is a sample item from the Degrees of Reading Power instrument.

> The Arctic is very far north. It is a cold land. It is icy. It is snowy. Winter is long. Winter is hard. Some plants can grow there. But nothing grows tall. Flowers are tiny. Grasses are low. Even the trees are __1__. They grow lying down on the ground. They can't grow tall. It is too cold.

1
a) wet b) short
c) dead d) gray
e) bare

Eskimos live in the Arctic. They know how to dress. They wear warm jackets and pants. These are made from animal skins. The fur is worn next to the body. They wear boots made from seal skins. These __2__ are important. They keep the Eskimo snug. They keep him dry.

2
a) clothes b) journeys
c) lamps d) jobs
e) homes

In the fall, the sun rises later each day. It sets earlier. Then the sun does not rise at all. It is dark all the time. There is hardly any __3__. This is the Arctic winter. It is cold. It is dark. It is hard to hunt. So the Eskimo must be ready. He must hunt before winter comes. He must find food. He must put it away. He must __4__ as much as he can. Then he will have food in winter.

3
a) wind b) water
c) light d) soil
e) industry

4
a) sleep b) save
c) move d) breathe
e) drink

Validity, Reliability, and Normative Data: The DRP is designed to measure the student's ability to comprehend text of a given level of readability (difficulty). Consequently, the difficulty or readability of each DRP passage is objectively determined using the Bormuth formula. (This formula takes into account such variables as word length, sentence length, and the proportion of difficult words in a passage.) The passages of the DRP are arranged so that each is one degree or step more difficult on the readability scale than the passage before it.

The difficulty of the deleted words and their options is held constant, and the test gets harder only with respect to the continuous text students must comprehend. Correctly answered test questions, therefore, indicate comprehension of text at increasing levels of difficulty and this ability measure is not confounded with the vocabulary difficulty of the response options.

Data for a pretest (January 1977)/posttest (June 1977) study ($n = 6{,}000$) were decomposed using the Rasch model. (If a subject took Form X in the pretest, Form Y was taken on the posttest six months later.) The pretest data for both forms scaled successfully; i.e., there was no item which did not fit the model. The linguistic factors (word frequency, sentence length, and word length) yielded a multiple of $R = .97$ with DRP passage difficulty for both forms. In a smaller study using matched subjects (by ability, socioeconomic status (SES), sex, and ethnicity), the difficulty of identical DRP items (sentences and response options) changed when the context was varied from easy to difficult. In the same study it was shown that having subjects read in-range material (prose within the student's comprehension ability) that contained correctly completed DRP items facilitated the performance on those same items when they were presented for completion in out-of-range prose (too difficult, relative to student ability). However, though there was a main effect for the manipulation, the only subjects who attained the mastery criterion for the passage (six out of seven correct) were those whose ability was close to the difficulty level of the difficult passages.

Data from another study (*n* = 1,500) showed that the most difficult level of text from which information can be acquired is the same for all subjects of given ability regardless of sex, ethnicity, or SES. No two-way or three-way interactions of ability with sex, race, or SES were significant. Therefore, DRP ability estimates do not have to be conditioned by any of these variables in order to predict information acquisition or gain on the six difficult cloze tasks.

Ordering Information: ED 170 712

Related Document:

University of the State of New York, State Education Department, Division of Educational Testing. *Degrees of Reading Power: Description of a New Kind of Reading Test and Its Related Technology.* Albany, N.Y. (ED 170 712)

Category: Reading

Title: Test of Oral Contrastive Stress

Author: Carole Kirchner Stice

Age Range: Upper elementary

Description of Instrument:

Purpose: To assess a reader's comprehension of items differing in contrastive stress.

Date of Construction: 1978

Physical Description: This test was devised as part of a research study to determine if a relationship existed between a reader's comprehension of contrastive stress and performance on a silent reading comprehension test. The test consisted of sixty-four items that were divided evenly among eight sentence types: declarative active, declarative passive, interrogative, and imperative, with each in a positive and negative form. The element of contrastive stress was placed on nouns functioning as the subject, verbs as the main predicate, nouns as the objects of prepositions, nouns as the direct objects, adjectives, and adverbs of time, place, and manner.

In order to reduce the problem of short-term memory, each subject was given a copy of the responses without the marked stress. The test required the subject to listen to a question followed by a sentence repeated three times with the element of contrastive stress being placed on a different word in each sentence to produce alternative meanings among the three sentences. The subjects were required to listen, read along on their copies, and select on their answer sheets the best rendition of the sentence to answer the question. The administration required approximately 45 minutes. An example is given below:

> If Arnold ate the cake rather than the pie, which way would you say this sentence?
>
> a. Arnold did not éat the pie.
> b. Árnold did not eat the pie.
> c. Arnold did not eat the píe.

Validity, Reliability, and Normative Data: The test was submitted for construct and content validation to a panel of experts consisting of two university professors of language arts, one professor of foreign language education, one professor of reading, and one professor of linguistics at Florida State University. The readability of each set of sentences (by type) was kept to approximately fourth grade or below. However, the results of a pilot study conducted on fourth, fifth, and sixth graders indicated that even though the vocabulary, length, and complexity of the test sentences were at fourth-grade level, the task required by the questions apparently required higher levels of cognition.

Reliability was assessed by the Kuder-Richardson Formula 20, and on the scores of 304 sixth graders, yielded a coefficient of .92.

Correlations between scores on the Test of Oral Contrastive Stress and The Comprehensive Test of Basic Skills, Level 2, Form Q, 1972 edition, were all significant at the .01 level.

Ordering Information: ED 236 635

Related Document:

Stice, Carole Kirchner. "The Relationship between Comprehension of Oral Contrastive Stress and Silent Reading Comprehension." *Research in the Teaching of English* (May 1978): 137–42. (EJ 186 574)

Category: Reading

Title: Test of Picture-Text Amalgams in Procedural Texts

Author: David Edey Stone

Age Range: Postsecondary

Description of Instrument:

Purpose: To assess how people read and comprehend information presented in picture-text amalgams in procedural texts.

Date of Construction: 1977

Physical Description: Although some research is available on the impact of illustrations on the comprehension of narrative texts, the opposite is true with regard to procedural texts. Since people are frequently confronted with completing tasks with the aid of illustrations, this test was considered to fill an important need.

The task chosen for use was selected from the Fishertechnik 100 model kit. The particular model used is the loading cart.

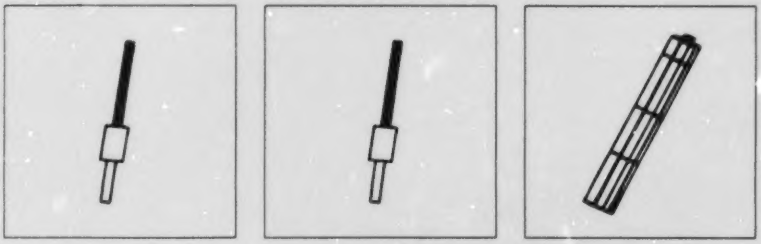
Slides containing various combinations of text information and illustrative information were prepared; subjects were assigned to one of four conditions and directed to follow the instructions presented on the slides. Through the use of videotapes, the following reading behaviors were observed: whether the reader was looking at a slide of text, at an illustration, or at the model parts; how long the reader's eyes remained in each of these three positions; how often the reader looked at each of these three positions; and the sequence in which the reader looked at them.

The first three steps of the text directions and the first three slides are presented below:

Text Directions

1. To form handle one: Insert a short rod through a clip so that the clip is in the middle of the rod.
2. To form handle two: Insert another short rod through another clip so that the clip is in the middle of the rod.
3. To form column one: Assemble one small block and three large blocks end to end.

Illustrations 1, 2, 3



Validity, Reliability, and Normative Data: A categorization system for tasks was developed and used as a guide for constructing the task directions. Four categories were identified: dimension, sequential constraints, number of substates, and the information present at the conclusion of the task concerning its preceding substates and their sequence.

The initial step in preparing the text directions involved having several people observe the performance of the task and record in writing what they observed. The task was performed very slowly so that ample time would be available for recording any aspects of the event deemed to be important.

The line drawings were prepared on a PDP 11 computer graphics system and were designed to closely approximate the pictorial directions in the manual of instructions for the model used. The graphics system used makes it possible to easily duplicate all or part of an existing drawing. This method of producing stimulus materials insured that the only differences in illustrations and text used would be those called for in the design of the experiment. For example, a line drawing representing a handle is constant regardless of whether it is shown singly or in combination with other parts.

On the basis of three pilot studies, appropriate changes were made until the task was considered appropriate for its purpose. In general, results of the study conducted with sixty subjects enrolled in an introductory psychology course showed that the addition of text resulted in a reduction in the amount of time spent looking at illustrations. The addition of text also resulted in more rapid and accurate task performance.

No reliability data for the subjects' completion of the task were reported.

Ordering Information: ED 236 665

Related Document:

Stone, David Edey. "Comprehension of Information in Picture-Text Amalgams in Procedural Texts." Ph.D. diss., Cornell University, 1977.

Category: Reading

Title: Orthographic Anomalies in a Silent Reading Task

Author: Michael Caley Strange

Age Range: Grades 5 and 6

Description of Instrument:

Purpose: To determine the degree to which good readers attend to the visual information presented to them. More specifically, to derive information about those parts of words that are most critical to the reader.

Whereas many studies have investigated the effects on readers of various orthographic factors within words, this study dealt with such factors in connected discourse and in a silent reading situation.

Date of Construction: 1976

Physical Description: Materials for the investigation were constructed through the following procedure. Seven passages, of 350 words each, were selected from the Scott, Foresman *Open Highways* basal reading series (1965). One passage was selected to be used to collect normal reading speed data and thus remained nonanomalized. The remaining six passages were systematically subjected to anomalization. Within each passage, the first 100 words were not anomalized. This procedure was completed in order to increase the likelihood that normal reading behavior would be reestablished after each trial. A given type of anomaly was then inserted in every tenth word for the remaining 250 words of the passage. This resulted in a package of materials containing seven passages—one nonanomalized passage and six anomalized passages. Each of the six anomalized passages represented a different combination of degree and position of anomaly.

The first anomaly type was beginning minor; that is, a minor anomaly was inserted into the beginning part (first third) of each of the twenty-five anomalized words. An anomaly was considered to be minor if it did not change the configuration of the word or was not a vowel-consonant substitution in the beginning or final position. The second anomaly type was beginning major. An anomaly was considered to be major if it did change the configuration of the word or was a vowel-consonant substitution in the beginning or final position. The third anomaly type was middle minor, a minor anomaly inserted in the middle third of a word. The fourth anomaly type was a middle major. The fifth anomaly type was a final minor, a minor anomaly inserted in the final third of a word. The sixth anomaly type was a final major. The following table presents a small section of prose, reflecting anomalies in context.

Samples of Anomalized Text

Type:	Example:
Beginning Minor	The reason was to de found in the window of the Park Square zporting goods store. It was a magnificent bicycle vith a sign in front of it that said:
Beginning Major	They returned to Peterson Park, lhe small but woodsy public park where Kirby and nis friends often played. Set free, the small, flop eared qog, still part puppy, went racing away.
Middle Minor	Poor old Miss Pecse would really have another fit if we hod a free-for-all again at one of her reheansals. She's just barely forgiven us for the last one.
Middle Major	"I'ye been out of town for a few years, but ndw I'm back to look things over and right away I sde something I don't like."
Final Minor	"Helle, son," said Mr. Maxwell, putting an arm around him az they walked toward the house. "Come on up and tall to me while I change clothes."
Final Major	"It doeg look more like a packing case that's been througn a train wreck. Still, if he were anybody elce, I'd be tempted to get with the Claypools and Burtons anq send him a petition."

The following directions were presented to the subjects on tape:

> Please read the following passages to yourself, one at a time. Each passage is two pages long. When you've finished each passage, look up at the person sitting with you. Read each passage at a comfortable speed. We are interested in your normal reading speed. We will be timing you, but please don't feel you have to

> rush. The person sitting with you will tell you when to begin each passage.

After the directions were presented, subjects were asked if they had any questions. Then they were reminded to read at a comfortable pace and not to be concerned about their speed. They were instructed to look up when they finished a passage. When the subjects indicated that they were ready to begin, the experimenter uncovered the passage and began timing. Timing was done with conventional stopwatches, evaluated to be accurate to the nearest tenth of a second. The watch was stopped when the subject looked up. The amount of time taken was recorded by the experimenter.

After reading each passage the subject was asked conventional recall questions about the passage, for example:

> Where was the boys' clubhouse? (in a vacant lot)
>
> What did the old man want them to do with the clubhouse? (tear it down)
>
> What did Kirby have to do if he wanted to win the bike? (write a poem)

The questions were not intended to assess the subject's complete comprehension of the material but only to determine that the subject had read the passage.

Validity, Reliability, and Normative Data: Construct validity may be claimed for the instrument. Word recognition is an accepted part of reading, although there is some question about which parts of words are most critical to the reader. The instrument more closely approximated a normal reading situation than other experiments in this area. The instrument allowed for the manipulation of the distinctive features of letters during the silent reading of connected discourse. On the basis of a pilot study with university undergraduate students, a decision was made to position the anomalies in every tenth word to make the task more suitable for upper elementary school pupils. On the basis of the Fry Readability Scale, the mean grade level for all test material was 4.5.

Results showed that anomalies in the middle of words were as disruptive to fluent reading as anomalies in the beginning and more disruptive than those at the end. Whether the anomalies were of a major or minor type was not an influencing factor.

No reliability data are reported.

Ordering Information: See related document.

Related Document:

Strange, Michael Caley. "The Effects of Orthographic Anomalies upon Good Readers' Performance." Ph.D. diss., University of Minnesota, 1976.

Category: Reading

Title: Walmsley CVC Patterns Test

Author: Sean A. Walmsley

Age Range: Primary

Description of Instrument:

Purpose: To provide a criterion-referenced measurement of CVC word patterns.

Date of Construction: 1975

Physical Description: In the construction of this test, an attempt was made to deal with two important perspectives of criterion-referenced tests: reading (especially procedures for defining a universe) and statistics (procedures for item generation and item analysis).

There are six subtests of 152 items each that were prepared as slides for projection. Each slide was made from 35mm mounts, with the test item pressed onto an acetate base with transfer lettering. A 10-point typeface in lower case was used, which produced a 1½-by-¾-inch image when projected onto a table viewer. Slides were mounted on a light orange acetate, and the table viewer was covered with a green acetate film, so that the contrast between the items (black) and the slide background would be decreased to acceptable and comfortable levels.

The items were presented to subjects on two separate occasions, usually separated by at least one day (seventy-six slides per session). The subject would be introduced briefly to the tester, made at ease, and told that he or she would be asked to read some three-letter words. Latency scores were also recorded.

Following is one item from the test that is approximately the size of the actual stimulus as seen by the subjects.

bat

Validity, Reliability, and Normative Data: In order to establish construct validity, great care was taken to define a CVC pattern and the universe to which it belonged. The following table indicates the criteria for the definition of a CVC.

Initial Consonant	*Medial Vowel*	*Final Consonant*
b, c[1], d, f, g[2], h, j, k, l, m, n, p, q[3], r, s, t, v, w[4], x[5], y, z.	a, e, i, o, u, y.	b, c[1], d, f, g[2], h[6], j, k, l, m, n, p, q, r[7], s[8], t, v, w[4], x, y[9], z.

[1] c has two pronunciations, /k/, /s/.

[2] g has two pronunciations, /g/, /γ/.

[3] *qVC* never occurs in the English language.

[4] *w* modifies the medial vowel *(a, o)* in the initial position; it has no sound as a final consonant-letter; when *w* is in the final position, the vowel is diphthongized.*

[5] *x* has the sound /ks/ but is usually pronounced /z/ when it occurs in the initial position of a syllable containing more than the letter *x* (i.e., *xerox*).

[6] *h* has no individual sound as a final consonant-letter; the combinations *ah, eh, oh, uh* are lengthened.

[7] *r* is sometimes (according to dialect) not pronounced as a consonant in the final position; here, the *r* may affect the vowel quality.

[8] *s* has two pronunciations, /s/, /z/.

[9] *y* does not usually (except in certain dialects) represent a consonant sound in the final position; when *y* is in the final position, the vowel is diphthongized.

Having generated all possible combinations of CVC patterns from these criteria, the combinations were then categorized according to real words/nonsense syllables; frequency as a word; frequency as a polysyllable; association value; and pronunciation value. In addition, they were categorized according to their pronunciation pattern (checked vowel, r-modified, etc.).

In order to provide for an adequate sample size of items, a sample ratio of 1:10 was adopted, yielding a total of 914 items which were then grouped into six subtests of 152 items. Thus, items appear in the test in numbers within categories that faithfully represent their proportion of the whole universe.

Concurrent validity was also established by correlating item scores from categories with reading ability, phonics knowledge, knowledge of real words, nonsense syllables, and r-w-y-words. All correlation coefficients were statistically significant.

---

*I include under diphthongs the following: /iy/; /ey/; /ay/; /aw/; /ow/; /uw/; /oy/.

The interrater reliability coefficient for two testers scoring the items for each of ten subjects was .995.

A general statement of the findings indicates that by the middle of second grade, the average second grader has a better than 60 percent understanding of the pronunciations of CVCs. By the middle of third grade, this understanding has improved to better than 65 percent, and by the middle of fourth grade, has improved further to better than 70 percent. At the same time, the speed with which a child recognizes and is able to pronounce a CVC has dropped from an average of over 3 seconds in second grade, to an average of under 2.5 seconds in third grade, and finally to an average of under 2 seconds in fourth grade.

Ordering Information: ED 236 633

Related Document:

Walmsley, Sean A. "The Criterion-Referenced Measurement of an Early Reading Behavior." Ph.D. diss., Harvard University, 1975.

Category: Reading

Title: Taxonomy of Reading Behaviors

Author: Jane White

Age Range: Primary through postsecondary

Description of Instrument:

Purpose: To describe the verbal protocols of subjects during in-process reading in a specialized setting.

Date of Construction: 1980

Physical Description: The taxonomy consists of six major categories, each of which is divided into subcategories. These are as follows:

1. Engagement in the Text—Measure of reader's involvement in the text.
    - 1.0 No response—Reader said nothing after a clause marker.
    - 1.1 Attempt to invalidate—Reader tried to invalidate the contract (of responding at each clausal marker) by saying only "I don't know" or similar response without paraphrasing or identifying problems, etc.
    - 1.2 Pseudo-engagements—Reader would begin to verbalize as asked, but stopped before the thought was completed, i.e., "I think '[illegible]ff' means. . . ."
    - 1.3 Maintenance of contract—Reader verbally responded at the clausal marker by paraphrasing, stating problems encountered, or any thoughts about the passage.
2. Response—Classification of the responses made by reader.
    - 2.1 Problem statement—Statement by reader indicated that something was interfering with comprehension.
    - 2.2 Capsulations—Statement summarized text information.
    - 2.3 Strategy usage—Statement indicated reader was predicting, confirming, integrating text information into personal schemata.
3. Unit of the Problem—Statement suggested that interference with comprehension was caused by a word/a sentence/the passage.
    - 3.1 With word—Statement indicated that problem was with a word in the passage.
    - 3.2 With sentence—Statement indicated that problem was with a sentence in the passage.
    - 3.3 With passage—Statement indicated that problem was at the passage level, no macrostructure instantiated to make the passage "fit."
4. Semantic problem—Statement suggested that meaning of words/concepts was disrupting comprehension.

4.1 Experience—Statement indicated there was not an experiential background for meaning with word/sentence/passage.

4.2 List schema—Statement indicated semantic problem was caused by the prior instantiation of a schema from the word list.

4.3 Previous reading of this text—Statement indicated reader was confused by the prior information given in the research passage.

5. Capsulations

5.1 Paraphrasing—Statement paraphrased text in reader's own words.

5.2 Repetition of text—Statement was an exact repeating of the text.

5.3 Relating text information to personal experience—Statement tied text information/understanding to reader's previous personal experience.

6. Strategies employed—Tactic or method a reader used to comprehend.

6.1 Predicting—Statement indicated reader predicted meaning from print.

6.1.1 Word level—Statement indicated reader predicted on the word level, i.e., "I think 'luff' means let the wind out of the sail."

6.1.2 Sentence level—Statement indicated reader predicted on a sentence level.

6.1.3 Passage level—Statement indicated reader predicted what passage meant or could mean.

6.2 Confirming—Reader accepted/rejected earlier predictions on the basis of further reading.

6.2.1 Word level—Statement indicated reader confirmed meaning of a word.

6.2.2 Sentence level—Statement indicated reader confirmed prediction of a sentence meaning.

6.2.3 Passage level—Statement confirmed prediction based on passage level macrostructure.

6.3 Integrating (comprehending)—Reader indicated that the reading made sense; it fit together.

6.3.1 Word level—Statement indicated integration occurred with a word.

6.3.2 Sentence level—Statement indicated integration occurred on sentence level.

6.3.3 Passage level—Statement indicated integration occurred on the passage, a note of completion, i.e., a sense of "Aha . . . this all fits."

In this taxonomy, the verbalizations had to be classified within "Engagement in the Text" (1). If the verbalization as a whole or any part of it was classified as "maintenance of contract" (1.3), then it had to be classified within "Response" (2). Once classified in the "Response" category as a "problem statement" (2.1), the finer descriptive categories within "Unit of the Problem"—"word," "sentence," or "passage" (3, 3.1, 3.2, 3.3)—and "Semantic Problem"—"experience," "list schema," or "previous reading of text" (4, 4.1, 4.2, 4.3)—could be used. If the verbalization or part of a verbalization was classified as "capsulations of text information" (2.2), then it could be classified within "Capsulations"—"paraphrasing," "repetition of text," or "relating text to personal experience" (5, 5.1, 5.2, 5.3). If the verbalization or part of it was classified as "strategy usage" (2.3) in the "Response" category (2), it could be further described in various subclassifications within "Strategies Employed"—"predicting," "confirming," "integrating" (6, 6.1, 6.2, 6.3)—with finer distinctions on the level of strategy usage—word, sentence, passage. A verbalization which matched a text clause was often several statements long. If so, the verbalizations had to be classified in as many ways as seemed necessary.

Once the descriptions of reading behavior were begun, the taxonomy became exceedingly flexible. A protocol could be classified in any number of categories within the taxonomy after an initial classification of "maintenance of contract" (1.3). A reader might be attempting to invalidate the contract (1.1) by saying "I don't know . . ." but then go ahead and maintain the contract (1.3) by mentioning a problem (2.1) on the word level (3.1) . . . "I don't know what 'luff' means" . . . and then predict (2.3) on the word level (6.1, 6.1.1), ". . . but I think it means sailing into the wind." The flexibility of the taxonomy is its strength. Because the data base used to develop the taxonomy consisted of the protocols of readers confronted with a difficult reading task, there was a supposition that there would be problem statements and statements describing the strategies being used to comprehend. These in fact did materialize and helped shape the taxonomy into its present form.

In the research study for which this taxonomy was devised, the author was also interested in determining differences in adaptation of one schema to a more appropriate one to fit the research text.

Consequently each subject was asked to create a story schema from a list of commonly used words (line, cloth, sheets, etc.) before silently reading the research passage. Due to the specific use of the list vocabulary in the research passage, the readers tended to predict a story schema which did not match the actual context, a story about a sailboat race. After reading each independent clause within the passage, each subject verbalized thoughts of either what the passage meant or problems he or she encountered while reading. The verbalizations or protocols were then transcribed and matched with the clause to which they referred. The effect of the schema orientation before the reading and the concomitant constraints on the reading process were investigated through the verbalizations.

An example of a passage used in the research study with the "schema" word list is given below.

Sailing Passage

The wind blew and everyone could see the lightning skim the rough surface and fork over the trees. Jim had to trim the sheets and head in.

After the worst of the storm had passed, Jim checked the pins and tightened the slipping cleats. While he was working Fred came by and said, "Terrible storm, isn't it?"

Jim replied, "It sure is. Have you ever seen one like this before?"

"Sure have," said Fred. "Once up at Evergreen, it blew so hard and was so cold I could hardly hang onto the sheets. I put her in irons and still wasn't in control."

"Well, I've got to wet sand the board," Fred continued. "Looks like the storm's blowing over so we can make the start. See you later."

When Jim heard the gun he yelled over his shoulder. "Let's rig up." They all headed out with the last trace of lightning showing across the bow.

As Jim's class approached the line, the blue flag came down and his hand tightened on the tiller. Jim fell off behind Fred's stern and crossed the line with the gun.

After that beautiful start, he knew he'd have clean air all the way. He rounded the windward mark and bore off on a run with the crew working feverishly to hoist the extra cloth. The effect was felt immediately as the speed picked up and they were able to plane. Both Fred and Jim swung wide at the leeward mark and Jim yelled at his crew to prepare to jibe as he cut inside. Jim and his crew moved as one on the tack up the windward leg. Maintaining a slight lead over Fred, Jim approached the finish line. As Jim came abeam of the finish line, the gun sounded to give Jim first place in the season's last sailboat race. He headed into the wind, allowing the main to luff. Friends grabbed his painter at the dock so he could receive his well earned trophy.

Another season was over. All the hard work had paid off. He'd finally beaten Fred in the season's last race.

Word list presented before sailing passage.

trim
sheets
pins
cleats
irons
board
class
line
clean
cloth
plane
tack
leg
painter

Validity, Reliability, and Normative Data: The taxonomy originates from a theoretical base similar to that used by Kintsch and Van Dijk (1978) in their identity of retelling categories. In order to insure content validity of the taxonomy, the protocols were divided by three reading specialists into classifications that seemed logical to describe the behavior exhibited in the protocols. The six major divisions of the taxonomy were mentioned by at least two of the specialists; five of the classifications were denoted by all three. The researcher added subclassifications to gain finer descriptive power.

In order to establish interrater reliability, the transcribed protocols were matched with the clause of the text to which each referred. Each protocol was then described as to the strategy employed or unit of text information referenced within that verbalization. Two independent raters classified four subjects' protocols to establish an interrater reliability with a correlation of .85 with the researcher.

Ordering Information: Full instrument reproduced above.

Related Documents:

Kintsch, W., and T. A. Van Dijk. "Toward a Model of Text Co[illegible]rehension and Production." *Psychological Review* 85 (September 1978): 36[illegible]-94.

White, Jane. "A Taxonomy of Reading Behaviors." Ph.D. diss., Southern Illinois University, Carbondale, Ill., 1980.

Category: Reading

Title: A Procedure to Evaluate Cognitive Requirements of Beginning Reading Materials

Author: Connie K. Williams

Age Range: Primary

Description of Instrument:

Purpose: The purpose of this instrument is to assess the cognitive requirement expected or implied in beginning reading materials and the instructional suggestions and to determine whether they are appropriate to the cognitive development of the children with whom they are to be used.

Physical Description: The instrument includes an introduction and directions for use. For consistency and so various materials can be compared, the rater is directed to examine the first reader from the basal series to be evaluated. It should be remembered that the selection of the first reader is for consistency when various series are to be compared and that this does not preclude the use of the procedure with other materials for beginning readers.

The main portion of the procedure consists of eleven items. Each item gives a brief explanation and examples of a certain aspect to be examined in the reader under consideration. A response from three alternatives that best describes the material is to be chosen for each item. A section is provided for totals to be determined. Three statements concerning the appropriateness of beginning reading materials for use with preoperational children are provided. After totals are analyzed, the statement that is most appropriate for the materials being reviewed is chosen.

The procedure was not designed to replace other rating devices and scales that can be used to evaluate other aspects of reading materials. Readability scales and other such procedures are recommended for use in conjunction with the procedure to evaluate cognitive requirements.

Following is a sample item:

> Are children required to hold in mind several things at one time? To obtain this information, analyze the activities suggested in the guided reading lessons as well as the workbook activities. An example of a workbook activity that requires holding more than one thing in mind is: "Color all the balloons with the long a's red, and color all the balloons with the short a's green except those followed by a silent e which should be colored blue." Be careful not to limit your examination to only this type of example.

_____ No activities found that require child to hold in mind several things at one time.

_____ Few activities (1 or 2) found that require child to hold in mind several things at one time.

_____ Several activities (3 or more) found that require child to hold in mind several things at one time.

Evidence:

In each section of the procedure, the first choice would indicate that the material is more desirable than if the third choice is checked. The middle choice is more desirable than the third but less desirable than the first.

After analyzing the totals, choose the statement which seems to be appropriate for the material which was reviewed.

_____ The materials are generally appropriate for use with preoperational children.

_____ The materials are somewhat appropriate for use with preoperational children.

_____ The materials are not appropriate for use with preoperational children.

Validity, Reliability, and Normative Data: Development of the Procedure to Evaluate Cognitive Requirements of Beginning Reading Materials involved four phases. First was an examination of the literature to serve as a basis for the development of the assessment procedure. Included was an examination of translations of the original works of Piaget as well as secondary works that provided interpretive information about his developmental theory and on which the instrument was based.

The second phase was also accomplished through the literature review. Research evidence and expert opinion were examined for identification of major principles for inclusion in the procedure to evaluate beginning reading materials in terms of cognitive requirements. After major principles were identified from the literature, validation of the principles was accomplished by judgments from a panel of experts in the fields of early childhood education, reading, and psychology. The panel members were selected from the faculty and participants of the Interdisciplinary Institute in Reading and Child Development (Waller 1977) or from suggestions of those participants. Upon consent to participate in this process, seven panel members were asked to rate each principle in terms of appropriateness for use with preoperational children and also for importance for inclusion in the procedure.

Panel members were asked to list additional activities, methods, or requirements that they would consider inappropriate for reading instruction with preoperational children; however, no additional principles were suggested. Based on the responses of the experts on the panel, coverage seemed to be complete, and content validity was supported. A weighted score was assigned to each response from panel members. A predetermined score was needed for inclusion of a principle in the exploratory version of the procedure. On this basis, the procedure was developed.

The third phase of the study involved the actual development of the procedure. The ratings and comments of the panel of experts guided the development. Consideration was also given to factors such as ease of use, time and training required, and objectivity of the procedure.

The fourth phase consisted of a pilot study of the procedure. Educators such as those who might participate in textbook selection decisions were asked to use the procedure and answer questions concerning its use. These educators represented five groups—first-grade teachers, elementary school principals, elementary school librarians, system-level supervisors, and college or university professors. Three people from each group were used for a total of fifteen participants in the pilot study.

So that comparisons could be made, all participants were instructed to use the same sample of instructional material. The text used was from a series based on synthetic phonics. Although the results of the pilot study were not totally consistent, a reasonable degree of consistency across raters was obtained.

The procedure was then used by the researcher with samples of three other types of reading programs—an analytic program, a linguistic program, and a program with literary emphasis. Use of the procedure with different approaches showed that the procedure discriminates between types of instructional material. Pilot study results were used to revise the procedure.

Ordering Information: ED 236 644

Related Document:

Waller, T. G. *Think First, Read Later! Piagetian Prerequisites for Reading.* Newark, Del.: International Reading Association, 1977. (ED 146 570)

Category: Reading

Title: Discourse-Analysis Based, Written, Multiple-Choice Post-Test for Comprehension Assessment of Expository Prose

Author: Petey Young

Age Range: Middle school/junior high

Description of Instrument:

Purpose: To assess immediate and delayed comprehension of expository prose.

Date of Construction: 1980

Physical Description: The instrument is a forty-item multiple-choice test composed of open-ended stems, each followed by four non-overlapping choices. Stems and choices are as short as possible. Control of the content of the items was based upon an analysis of the accompanying 1,300-word passage on the Kalahari Desert using Turner and Greene's (1977) directions for following Kintsch's (1974) system of discourse analysis. The analysis established whether bits of information were comparatively general and interrelated or detailed and isolated. After the analysis had been collapsed to three levels of generality by including the more detailed levels in the third level, equal numbers of items were directed at each of the three levels. The validity of the semantic match between test item and mother passage was established by comparing an analysis of the items with that of the passage.

Validity, Reliability, and Normative Data: The instrument was revised and validated by comparisons between the performances of two groups of students: one group took the test without reading the passage upon which the test was based; the other read the passage and had it available while taking the test. Revision consisted of deleting items answered correctly by a significant number of the first group and incorrectly by a significant number of the second group. After revision, one-way analyses of variance showed the differences between performances of the groups to be significant at a $p < .001$ level. The Hoyt reliability value for the group that read the passage was high, 0.89. When the instrument was used to gather data, treatment groups outperformed control groups at a $p < .001$ level of significance on both the immediate and delayed administration in two studies involving a total of 240 seventh graders.

Ordering Information: ED 236 656

Related Documents:

Kintsch, W. *The Representation of Meaning in Memory.* Hillsdale, N. J.: Lawrence Erlbaum Associates, 1974.

Turner, A., and E. Greene. *The Construction and Use of a Propositional Text Base*. Technical Report No. 63. Boulder: University of Colorado, Institute for the Study of Intellectual Behavior, 1977.

Young, P. *Effect of Inference-Making Aids on Poor Readers' Comprehension.* Technical Report No. 555. Madison, Wis.: Wisconsin Research and Development Center for Individualized Schooling 1980. (ED 193 607)

Category: Reading

Title: An Instrument for Assessing Comprehension through the Ability to Recognize Verbatim Phrases from Previously Read Expository Prose

Author: Petey Young

Age Range: Middle school/junior high

Description of Instrument:

Purpose: To be used along with a multiple-choice instrument for further assessment of the immediate and delayed comprehension of a 1,300-word expository passage.

Date of Construction: 1980

Physical Description: The instrument is a thirty-five-item posttest, divided into five clusters, each composed of six to eight lettered phrases. Throughout the clusters, twenty verbatim phrases from a 1,300-word passage on the Kalahari Desert were scattered randomly. The other fifteen phrases did not imitate the wording of the passage nor were they true according to the passage or to further knowledge of the topic. Instructions were to circle the letters of any phrases that occurred in the text. The selection of verbatim phrases was based on an analysis of the passage made by using Turner and Greene's (1977) directions for following Kintsch's (1974) system of discourse analysis. The analysis was made to yield three increasingly detailed levels of information by collapsing any more detailed levels to the third level. An equal number of verbatim phrases was then chosen from each level. This way, general, interrelated information was emphasized but specific, isolated bits of information were not completely neglected.

Validity, Reliability, and Normative Data: To ascertain whether or not the content or style of the wording signaled the correct responses, the instrument was administered to a group that did not read the passage but were told it was about the Kalahari Desert and given the task of guessing which phrases were actually in the story. To validate and revise the instrument, the performances of this group were compared to the performances of a group that read the passage and then took the test. The second group outperformed the first at a $p < .001$ level of significance. The test was revised by deleting items answered correctly by a significant number of the first group and incorrectly by a significant number of the second group. The Hoyt reliability of the group that read the passage was high, 0.80. When the instrument was used along with a forty-item multiple-choice test to gather data, treatment groups outperformed control groups on the phrase-recognition test at a $p < .001$ level of significance on both the immediate and delayed administrations in two studies involving 240 seventh graders. Pearson product-moment correlations found between the phrase-recognition test and the multiple-choice test scores ranged from .27 to .68 in the treatment and control groups. The moderate positive correlations indicate the two tests measured related, but not identical, abilities.

Ordering Information: ED 236 655

Related Documents:

Kintsch, W. *The Representation of Meaning in Memory*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1974.

Turner, A., and E. Greene. *The Construction and Use of a Propositional Text Base*. Technical Report No. 63. Boulder: University of Colorado, Institute for the Study of Intellectual Behavior, 1977.

Young, P. *Effect of Inference-Making Aids on Poor Readers' Comprehension*. Technical Report No. 555. Madison, Wis.: Wisconsin Research and Development Center for Individualized Schooling, 1980. (ED 193 607)

Teacher Knowledge/Attitudes

Category: Teacher Knowledge/Attitudes

Title: Categories for Observing Language Arts Instruction (COLAI)

Author: Julianna G. Benterud

Age Range: Any instructional situation

Description of Instrument:

Purpose: To study individual use of learning time, with particular focus on time spent in reading during the time scheduled for language arts in a natural classroom setting.

Date of Construction: 1983

Physical Description: A coding sheet was devised on which the standardization of the moment of observation and of the interval between observations for each pupil was achieved by premarking the coding sheet in multi-second intervals. The study for which the coding sheet was devised involved four subjects, each of whom was observed at the beginning of a 30-second interval. That is, if subject one was observed at 9:00:00, subjects two, three, and four would be observed at 9:00:30, 9:01:00, and 9:01:30 respectively, with the rotation again beginning for subject one, at 9:02:00. The behavior noted during the particular observation time was coded in the appropriate column on the coding sheet.

The coding sheet consisted of nine major columns. These columns with their subcategories are as follows:

1. Engagement
   a. Definitely engaged (D)
   b. Definitely not engaged (DN)
   c. Can't tell (CT)
2. Area of Language Arts
   a. Listening (L)
   b. Speaking (S)
   c. Viewing (V)
   d. Writing (W)
   e. Oral reading (Ro)
   f. Silent reading (Rs)
3. Instructional setting
   a. Whole group (WG)
   b. Small group (SG)
   c. Individual (I)
4. Partner
   a. Teacher (T)
   b. Pupil or pupils (P)
5. Source of content
   a. Worksheet (WS)
   b. Workbook (WB)

  - c. Basal reader (BR)
  - d. Library book (LB)
  - e. Chart (Ch)
  - f. Blackboard (BB)
  - g. Other (O)
6. Type of unit
   - a. Pictures (P)
   - b. Letters (L)
   - c. Isolated words (IW)
   - d. Isolated sentences (IS)
   - e. Connected discourse (CD)
7. Assigned or chosen task
   - a. Assigned general (AG)
   - b. Assigned specific (AS)
   - c. Chosen (Ch)
8. Rate of success
   - a. High success (H)
   - b. Medium success (M)
   - c. Low succcess (L)
9. Other activities
   - a. Transition (T)
   - b. Wait (W)

Validity, Reliability, and Normative Data: The observation categories are based on the underlying construct "Academic Learning Time" that originated in the Beginning Teacher Evaluation Study conducted over the six years between 1972 and 1978 by the Far West Laboratory for Educational Research and Development (Berliner 1979). ALT, defined as the amount of time a student spends attending to academic tasks while also performing at a high rate of success, was shown to be a valid indicator of student learning on the basis of the high correlation between ALT ratings and achievement test scores. In addition, the ALT has the advantage over achievement tests of providing a means of assessing ongoing, in-process learning.

Content validity was obtained for the COLAI by utilizing the overall structure of the coding system developed for the Beginning Teacher Evaluation Study. Modifications were made on the basis of a pilot study conducted over a seven-week period in five first-grade classrooms.

Ecological validity may be claimed on the basis of the instrument being used in a natural class setting to collect data during the 600 minutes allocated to language arts instruction during a single week.

Interrater reliability based on the ratings of two independent observers for each of the nine major categories prior to and during the study resulted in coefficients of agreement ranging from .90 to .98 for a 70-minute observation period prior to the study, and between .94 and .98 for a 90-minute observation period during the study.

Ordering Information: ED 236 632

Related Documents:

Benterud, J. G. "From First-Graders' Use of Reading Time During the Language Arts Period in Their Natural Classroom Setting." Master's thesis, University of Alberta, Edmonton, Alberta, Canada, 1983.

Berliner, D. C. "Tempus Educare." In *Research on Teaching: Concepts, Findings and Implications,* edited by P. L. Peterson and H. J. Walberg. Berkeley, Calif.: McCutchan Publishing Corp., 1979.

Category: Teacher Knowledge/Attitudes

Title: The Chin Inventory on Content Area Reading Instruction

Author: Beverly Ann Chin

Age Range: Adult

Description of Instrument:

Purpose: To assess attitudes and perceptions of competency in teaching reading in the content areas held by preservice and inservice teachers of elementary and secondary content area subjects.

Date of Construction: 1975

Physical Description: Teachers respond to twenty-seven specific skill statements on a 5-point Likert-type scale. Each statement represents and illustrates the integration of reading instruction with content instruction. The Chin Inventory has two independent sections. Part A assesses teachers' attitudes toward content area reading instruction. Part B assesses teachers' perceptions of competency in content area reading instruction. Sample items from each section follow:

PART A

Below is a set of statements dealing with reading instruction in the content areas. Rate the *IMPORTANCE* of each statement to you as a teacher in YOUR CONTENT AREA by responding to the following 5-point scale:

Very Important 1 2 3 4 5 Not Important

Mark your responses by circling your choice on the scale after each statement.

To me, incorporating into my assignments instruction on how to read regular classroom materials is

Very Important 1 2 3 4 5 Not Important

PART B

Below is a set of skills dealing with reading instruction in the content areas. Rate *YOUR QUALIFICATION* to perform each skill in YOUR CONTENT AREA by responding to the following 5-point scale.

Very Qualified 1 2 3 4 5 Not Qualified

Mark your response by circling your choice on the scale after each skill question.

How Qualified Are You

To assist students in setting a definite purpose for reading assigned materials?

Very Qualified 1 2 3 4 5 Not Qualified

Each section of the Chin Inventory should be scored separately. To assess teachers' attitudes toward teaching reading in the content area, Part A of the Chin Inventory should be given. Teachers indicate the degree and direction of their attitude toward content area reading instruction by rating the importance of each of the twenty-seven statements on the 5-point scale. The number selected on the attitude scale becomes the weighted response assigned to that item. The sum of the twenty-seven weighted responses is the teacher's attitude score, with a low total score indicating a more positive attitude toward content area reading instruction than a high total score. A total attitude score can range from 27 (1 × 27) to 135 (5 × 27), and any score below 81 (3 × 27) reflecting an attitude on the positive side of the scale.

To assess teachers' perceptions of competency in teaching reading in the content area, Part B of the Chin Inventory should be given. Teachers indicate the degree and direction of their competency in content area reading instruction by rating their qualification to perform each of the skills on a 5-point scale. The number selected on the competency scale becomes the weighted response assigned to that item. The sum of the twenty-seven weighted responses is the teacher's perception of the competency score, with a low total score indicating a more positive perception of competency in content area reading instruction than a high total score. A total perception of competency score can range from 27 (1 × 27) to 135 (5 × 27), with any score below 81 (3 × 27) reflecting a perception of competency on the positive side of the scale.

Validity, Reliability, and Normative Data: The content validity of the Chin Inventory was determined by a thorough review of the professional literature on content area reading instruction. Based on this survey, specific competency statements were created to exemplify the integration of reading instruction into content instruction. These competency statements were submitted to a panel of content area teachers and reading experts who evaluated each statement. The statements that were judged to be important competencies in content area reading instruction, applicable to teachers in all content areas, became items of the instrument. In addition, the total set of skill statements presented an operationally defined, competency-based view of the concept of content area reading instruction.

The reliability of the Chin Inventory was determined by the split-half method. Reliability coefficients of .92 for the attitude measure (Part A) and .97 for the perception of competency measure (Part B) were obtained using the Spearman-Brown prophecy formula on a sample population ($n = 224$). The coefficients obtained for each of the subgroups in the sample population ranged from .86 to .98, all of which were highly reliable for group measurement purposes.

Ordering Information: ED 236 657

Related Documents:

Evans, Carolyn. "Attitudes of Secondary Teachers toward Teaching Reading and Their Perceptions of Their Competencies in Reading Instruction." Ph.D. diss., Mississippi State University, 1977.

Flanagan, Beverly Ann Chin. "A Competency-Based Assessment of Secondary Teachers' Attitudes and Perceptions of Qualification in Content Area Reading Instruction." Ph.D. diss., University of Oregon, 1975.

Kozey, Louisa. "A Study of Secondary Teachers' Perceptions of Importance, Qualifications, and Implementation of Content Area Reading Instruction." Ph.D. diss., University of Oregon, 1980.

Category: Teacher Knowledge/ Attitudes

Title: The DeFord Theoretical Orientation to Reading Profile (TORP)

Author: Diane E. DeFord

Age Range: Adult

Description of Instrument:

Purpose: To differentiate preservice and inservice teachers according to their theoretical orientation to reading.

Date of Construction: 1979

Physical Description: The instrument consists of twenty-eight items reflecting practices and beliefs about reading instruction. Each item is responded to on a 5-point Likert scale indicating strength of agreement with the statement (strongly agree to strongly disagree). The instrument necessitates a forced response, and, although an individual's belief system may fall anywhere along the continuum (as indicated by the scale), there are three major clusters of orientations which may be called Phonics (smaller-than-word emphasis), Skills (whole words with multiple skills for dealing with this unit), and Whole Language (larger-than-word segments). The resulting score is a general indicator of the respondent's orientation. A score within the lower range (0–65) would indicate an orientation toward phonics, within the middle range (65–110) toward skills, and within the high range (110–140) toward whole language. The instrument takes 20 to 30 minutes to administer. Sample items:

A child needs to be able to verbalize the rules of phonics in order to assure proficiency in processing new words.

1 2 3 4 5
SA SD

It is a good practice to correct a child as soon as an oral reading mistake is made.

1 2 3 4 5
SA SD

Validity, Reliability, and Normative Data: The instrument was subjected to a pilot study and was reviewed by professionals in the field prior to the final rewrite. Validity and reliability data were obtained by securing data about the construct (teachers' belief patterns) from a number of sources. The first was to administer the TORP to ninety teachers identified by educators familiar with their teaching (thirty phonics, thirty skills, thirty whole language). Teacher responses by group did reflect differences in means supporting different profiles for each group, and the readability coefficient was reported as a .98 across the three different groups. Factor analysis produced one factor that accounted for 94.5 percent of the variance, from which it can be inferred that the TORP is a one-factor test measuring instruction in reading characterized by a continuum from isolation to integration of language.

The second step involved teacher observations ($n = 14$) by trained observers ($n = 4$). After observing a teacher, each observer responded to the TORP so as to reflect the teacher's orientation. The teachers responded to the instrument also. The teacher/observer overall correlation with a Spearman Rho rank order correlation was .859 ($p < .001$). This correlation was based upon a rank ordering of teacher total score with observer total score across the fourteen teachers.

Step three utilized a pretest/posttest analysis from an undergraduate methods course in reading and language arts instruction ($n = 20$). A t-value of 15.05 for twenty-eight degrees of freedom ($p < .01$, two-tailed) indicated a significant change toward the instructor's theoretical orientation.

The final step in the validation process required judges ($n = 3$) who were in charge of teachers in practicum settings to respond to the TORP in three different ways. They were asked to reflect the responses of a phonics, then a skills, and, finally, a whole-language orientation. The comparison of the judges' responses resulted in a Kendall's W of .83 and a chi-square of 205.65 with 83 degrees of freedom, which was significant beyond the .001 level.

Ordering Information: ED 236 661

Related Document:

DeFord, Diane E. "A Validation Study of an Instrument to Determine a Teacher's Theoretical Orientation to Reading Instruction." Ph.D. diss., Indiana University, 1979.

Category: Teacher Knowledge/Attitudes

Titles: 1. Knowledge of Content Area Reading Skills
2. Situations Survey: Teaching Reading in Content Areas
3. Statements Survey: Teaching Reading in Content Areas

Authors: Mary M. Dupuis
Eunice N. Askov
Joyce W. Lee
Carlotta Joyner Young (2,3)
Jeffrey McLoughlin (1)

Age Range: Adult

Description of Instrument:

Purpose: To assess knowledge and attitudes about integrating reading skills with content area instruction held by preservice and inservice teachers of content area subjects, grades 4–12.

Date of Construction: 1. 1979
2. 1976
3. 1976

Physical Description: Knowledge of Content Area Reading Skills is a criterion-referenced measure consisting of thirty-four multiple-choice items designed to test knowledge of basic materials and methods for teaching reading in various content areas. Respondents have five choices for each item, with the fifth always reading "I honestly don't know," in order to reduce haphazard guessing. Sample:

> The primary task of the content area teacher with respect to diagnostic teaching of reading is to:
>
> a. foster the transfer of basic reading skills to content area materials.
> b. develop positive attitudes toward content subjects.
> c. provide phonics instruction to those students who need it.
> d. assess student performance in relation to graph norms.
> e. I honestly don't know.

Situations Survey: Teaching Reading in Content Areas, which utilizes the semantic differential technique, consists of twelve items with five sets of bipolar adjectives (such as *useful-useless*) to be rated for each item. Each item consists of a classroom situation that a content area teacher might face and a possible diagnostic-prescriptive plan the teacher might follow in the situation described. Two questions using the semantic differential format and included as part of the Situations Survey yield two additional scores used in assessing the effects of an inservice program. The first of these scores, the Feasibility Score, is obtained from teacher ratings of the bipolar adjectives *feasible-not feasible* after each of the twelve items on the Situations Survey. The other score is a self-report measure consisting of teacher ratings of the bipolar adjectives *skilled-not skilled* after each of the twelve items on the Situations Survey. This Perceived Skill score was designed to measure a teacher's confidence in implementing the stated diagnostic-prescriptive plan. Sample:

> *SITUATION:* An English teacher is preparing to teach a short story from the anthology suggested in the curriculum guide.
>
> *PLAN:* The teacher plans to assign those who are competent readers to read the story on their own and engage in several individualized assignments. The less competent readers will read the story in a guided reading lesson during which the teacher will provide considerable help in vocabulary concept development, and comprehension.
>
> practical ___:___:___:___:___:___:___ impractical
> ineffective ___:___:___:___:___:___:___ effective
> inefficient ___:___:___:___:___:___:___ efficient
> useful ___:___:___:___:___:___:___ useless
> desirable ___:___:___:___:___:___:___ undesirable
>
> On the basis of your classroom experience, how feasible would you say the above plan is?
>
> feasible ___:___:___:___:___:___:___ not feasible
>
> How skilled are you at this time for executing a plan like the one described above?
>
> skilled ___:___:___:___:___:___:___ unskilled

Statements Survey: Teaching Reading in Content Areas is a 20-item Likert scale measuring teacher attitudes toward incorporating reading instruction in the content areas. Sample:

> It is important that teachers be competent in assessing the general reading levels of students.

(a)	(b)	(c)	(d)	(e)
Strongly Disagree	Slightly Disagree	Not Sure	Slightly Agree	Strongly Agree

Validity, Reliability, and Normative Data: The Knowledge of Content Area Reading Skills has an estimated test-retest reliability of .82. Its validity is based on matching each item with competencies included in the Content Area Reading Project and by comparison of topics tested with those in textbooks and other publications related to content area reading.

Estimated reliabilities are .96 (coefficient alpha) for the Situations Survey, .91 (test-retest) for the Perceived Skill Score, .75 (coefficient alpha) for the Feasibility Score, and .84 (coefficient alpha) for the Statements Survey. The Situations Survey, the Statements Survey, and an earlier version of the Knowledge Test have been used with inservice teachers since 1976 as part of the Content Area Reading Project of the Pennsylvania Department of Education.

Ordering Information: ED 155 666

Related Document:

Dupuis, Mary M., Eunice N. Askov, and Joyce W. Lee. "Changing Attitudes toward Content Area Reading: The Content Area Reading Project." *Journal of Educational Research* 73 (November/December 1979): 65-74. (EJ 216 888)

Category: Teacher Knowledge/Attitudes

Title: Criterion-Referenced Test for the Assessment of Reading and Writing Skills of Professional Educators

Authors: Mary M. Dupuis
Sandra L. Snyder

Age Range: Adult

Description of Instrument:

Purpose: To assess the reading and writing skills of professional educators.

Date of Construction: 1979

Physical Description: The criterion-referenced test, developed as a measure of professional reading skill of preservice and inservice teachers, requires the reading of an article in a professional journal [Barry J. Wilson and Donald W. Schmits, "What's New in Ability Grouping?" *Phi Delta Kappan* 60 (April 1978): 535–36]. Teachers are assessed in four areas reflecting Barrett's levels of questioning, yielding these four sub-scores plus a total score: ability to understand the professional vocabulary used in the selection, ability to answer literal-level comprehension questions, ability to answer inferential-level comprehension questions, and ability to interpret information found in tables. The writing sample is a response to one of two evaluative questions related to the same reading selection. The responses are scored with holistic procedures, using a rating system based on the CEEB 4–1 scale, on which 4 is the highest and 1 the lowest score. Although the test is untimed, the entire assessment, including reading the selection, completing the twenty-three-item multiple-choice test, and writing the assignment, can be completed in 60–75 minutes by most subjects. Sample items:

Comprehension
1. In the study of ability grouping conducted by the authors, most teachers familiar with research results:
   a. preferred heterogeneous grouping.
   b. felt ability grouping contributes to division among social classes.
   c. favored ability grouping.
   d. believed ability grouping hinders learning.

Vocabulary
15. A *generic* term is:
   a. prevalent.
   b. educational.
   c. classic.
   d. collective.

Data Interpretation (from a table)

21. Which ability group benefits most from homogeneous grouping?
    a. low
    b. average
    c. high
    d. impossible to tell from the data given

Writing Exercise

1. What does the information contained in this article suggest about the influence of research results on instruction? In your opinion, is this an accurate representation of the application of research findings to instructional practices in general? Suggest ways that research could have a greater influence on instruction.

Validity, Reliability, and Normative Data: Validity was established by a comparison of the objectives to standard reading assessments, both in terms of skills assessed and in procedures followed. The use of a professional article on a timely topic in a reputable journal enhances validity since teachers are being assessed using reading material similar to that which they will be reading as part of their continuing professional growth. No reliability data are available, though authors plan a test-retest procedure for the reading assessment. Interrater reliability for the scoring of the writing sample was .71 for a first assessment, and .93 for a second assessment. Normative data were collected for forty-three preservice teachers.

Ordering Information: ED 236 643

Related Documents:

Barrett, Thomas C. "Taxonomy of Reading Comprehension." *Reading 360 Monograph*. Lexington, Mass.: Ginn and Company, 1972.

Fagan, Edward R., Mary M. Dupuis, and Sandra L. Snyder. "Basic Skills of Prospective Secondary Teachers." *English Record* 32 (Spring 1981): 14–18.

Category: Teacher Knowledge/Attitudes

Title: Emig-King Writing Attitude Scale for Teachers

Authors: Janet Emig
Barbara King

Age Range: Adult

Description of Instrument:

Purpose: To measure attitudes toward writing and changes in attitudes toward writing held by preservice and inservice teachers.

Date of Construction: 1979

Physical Description: This instrument is a revision of the teacher version of the "Emig Writing Attitude Scale" that was constructed in 1977 for the New Jersey Writing Project. The revised scale contains fifty statements representing three categories: preference for writing, perception of writing, and process of writing. Approximately 30 minutes are required for the administration of the scale that asks teachers to circle one of 5 points ranging from "almost always" to "almost never." Sample items:

> Preference for Writing
>
> I accept positions in groups that involve writing.
>
> Almost always Often Sometimes Seldom Almost never
>
> Perception of Writing
>
> Studying grammar formally helps students improve their writing.
>
> Almost always Often Sometimes Seldom Almost never
>
> Process of Writing
>
> I revise what I write.
>
> Almost always Often Sometimes Seldom Almost never

Validity, Reliability, and Normative Data: The teacher version of the original Emig Writing Attitude Scale was administered to twenty-five teacher-participants in the 1977 New Jersey Writing Project. The revised Emig-King Writing Attitude Scale for Teachers was administered to fifty-three teachers who participated in the 1979 Summer Institute of the New Jersey Writing Project. Items on the revised scale were submitted to graduate students in English education and secondary teachers of English in order to establish content validity. Suggestions made by these experts were considered and revisions made. The Cronbach alpha reliability was .788.

Ordering Information: ED 236 629

Related Document:

King, Barbara. "Two Modes of Analyzing Teacher and Student Attitudes toward Writing: The Emig Attitude Scale and the King Construct Scale." Ph.D. diss., Rutgers University, 1979.

Category: Teacher Knowledge / Attitudes

Title: Gary-Brown Writing Opinionnaire for College Instructors

Authors: Melvin Gary
Sandra Brown

Age Range: Adult

Description of Instrument:

Purpose: To assess teachers' attitudes toward teaching writing, evaluating writing, and teaching course content through writing.

Date of Construction: 1981

Physical Description: Part I of this two-part measure is a Likert-type scale containing thirty-five items representing the instrument's three purposes. Responses range across 5 points from "strongly agree" to "strongly disagree." Sample items:

	Strongly Agree	*Agree*	*No Opinion*	*Disagree*	*Strongly Disagree*
Evaluating Writing					
3. Specific penalties in grades should be assigned for mechanical errors in term papers—for example, points should be deducted for each grammatical error.	☐	☐	☐	☐	☐
Knowledge of the Writing Process					
16. Students should engage in a prewriting process prior to writing the first draft of their compositions.	☐	☐	☐	☐	☐
Teaching Course Content through Writing					
27. Writing is a primary mode of learning as well as a tool for measuring learning.	☐	☐	☐	☐	☐

Part II asks for professional opinions in twelve categories, including extent of marginal comment on student papers, nature of writing assigments, peer evaluation, percentage of college students with severe writing problems, most common student writing errors, rewriting, and recommendations for improvement of instruction. Approximately 30 minutes are needed for administration of the total instrument.

Validity, Reliability, and Normative Data: The scale is in use by 227 faculty members at Livingston College, Rutgers University. No reliability or normative data have thus far been reported. The present scale was constructed after suggested revisions were made by college administrators, professors of English, and professors of English education. The suggestions of these experts contributed to the content validity of the scale.

Ordering Information: ED 236 660

Related Document:

Commission on Composition, National Council of Teachers of English. *Composition Opinionnaire: The Student's Right to Write.* Urbana, Ill.: National Council of Teachers of English and the ERIC Clearinghouse on Reading and Communication Skills, 1972. (ED 091 729)

Category: Teacher Knowledge/Attitudes

Title: A Survey of Methods and Materials

Authors: Candida Gillis
Lois Rosen
Wendy Neininger

Age Range: Adult

Description of Instrument:

Purpose: To assess and describe the nature and frequency of teaching methods, activities, and materials used in secondary school English courses.

Date of Construction: 1976

Physical Description: The survey consists of fourteen questions related to teaching objectives, activities, and materials, each followed by a list of possible responses ranging in length from seven to forty-three items. The directions ask teachers to select a course, and to circle for each of the 256 total items the frequency with which he or she used a particular activity, material, or objective in that course (0 = never, 1 = one or two times during the term, 3 = frequently). Questions cover the concepts taught to improve students' abilities in English; the activities used to motivate students to read literature and to write; the techniques for teaching writing, literature, or reading; the kinds of writing students are assigned and the nature of reading materials that are used in class; the reasons for selecting reading materials; the range of oral activities; the nature of any non-written composing; the methods of evaluating student progress; and the teachers' purposes for having students read, write, and speak in classrooms. The directions also ask teachers to identify the nature of the class for which they are completing the survey, including the general ability of the students. These variables are used to analyze data from large samples of teachers.

Sample:

Did you use any of the following to motivate students to write?

Movies or television	0	1	3
Slides, filmstrips, or pictures	0	1	3
Literature	0	1	3
Music	0	1	3
Guest speakers	0	1	3
Whole class discussions	0	1	3
Small group discussions	0	1	3

Creative dramatics	0	1	3
Field trips	0	1	3
Teacher talk or lecture	0	1	3
Student-produced materials—writing, reports, projects, etc.	0	1	3
Other ______________________	0	1	3

Validity, Reliability, and Normative Data: The items used for each question were chosen to present teachers with a wide range of representative activities, concepts, and materials that might be used in a variety of class settings, for a variety of student abilities, and by teachers favoring different approaches. Items were field tested on samples of teachers, including teachers whose students were of different ages and abilities and whose teaching encompassed different areas of the English curriculum to make sure that they accurately reflected teachers' self-perceived practices. Additionally, the survey was reviewed by members of the Standing Committee on Research of the National Council of Teachers of English. Data on reliability are not available. Normative data exist for 595 teachers, including those who filled out the survey when it was published in the *English Journal* as a 1977 readership survey.

Ordering Information: ED 236 641

Related Document:

Gillis, Candida. "The English Classroom 1977: A Report on the *EJ* Readership Survey." *English Journal* 6 (September 1977): 20–26. (EJ 167 565)

Category: Teacher Knowledge/Attitudes

Title: Mastery Assessment of Basic Reading Concepts (MABRC)

Author: Robert A. Pavlik

Age Range: Adult

Description of Instrument:

Purpose: To determine if elementary school teachers have mastered reading concepts most emphasized in undergraduate courses, where they mastered these concepts, or why they have not mastered the concepts.

Date of Construction: 1974

Physical Description: This measure includes fifty-one items in multiple-choice format. The items cover "basic reading concepts," defined to include: nature of the reading process, the developmental reading program, reading readiness, approaches to beginning reading instruction, reading skills, reading assessment, and reading instructional practices. Illustrated below is a sample item and its possible responses:

Column A	*Test Item*	*Column B*
A ______	A. The reading readiness factor which the classroom teacher can do *least* to improve is ___ 1. auditory discrimination. ___ 2. experiential background. ___ 3. language facility. ___ 4. socio-economic status.	B ______

Possible Responses for Column A:	Possible Responses for Column B:
1 — I mastered this concept in undergraduate reading methods courses.	1 — I never encountered this concept.
2 — I mastered this concept in my teaching experience.	2 — This concept is confusing and abstract.
3 — I mastered this concept somewhere else. State where on blank provided.	3 — My undergraduate preparation and teaching conflict on this concept.
4 — I am not sure where I mastered this concept.	4 — I encountered this concept, but I have forgotten my understanding of it.

Validity, Reliability, and Normative Data: The reading education professors at the University of Northern Colorado wanted to follow up on their former students and this instrument was designed to serve that purpose. It was additionally a part of an evaluation of the undergraduate reading program at the University of Northern Colorado. As such, the instrument and the data it gathers are idiosyncratic to the professors, and the concepts they emphasized in their undergraduate reading education courses. An item analysis was conducted on the original version of this measure, and a revision was developed based on the results. As part of the author's doctoral dissertation, the instrument was used with 346 elementary teacher education program graduates.

Ordering Information: ED 236 649

Related Document:

Pavlik, Robert Andrew. "An Assessment of Undergraduate Preparation for the Teaching of Reading in the Elementary School." Ph.D. diss., University of Northern Colorado, 1974.

Category: Teacher Knowledge/Attitudes

Title: Teacher Attitudes toward Composition Instruction

Authors: Brian F. Schuessler
Anne Ruggles Gere
Robert D. Abbott

Age Range: Adult

Description of Instrument:

Purpose: To assess teacher attitudes toward four components of composition instruction: (1) standard English, (2) linguistic maturity, (3) defining and evaluating writing tasks, and (4) student self-expression in the instruction of written composition.

Date of Construction: 1980

Physical Description: Four ten-item scales measure teacher attitudes toward four areas of composition instruction. Responses to items are recorded on a 5-point Likert-type scale anchored by: (1) strongly disagree, (2) disagree, (3) no opinion, (4) agree, and (5) strongly agree. Following are titles for the four scales and sample items.

The Importance of Standard English in the Instruction of Written Composition

1. In order to avoid errors in sentence structure, weak students should be encouraged to write only short, simple sentences.
2. High school students should be discouraged from using figurative language because their efforts at metaphor so often produce only clichés.

The Importance of Defining and Evaluating Writing Tasks in the Instruction of Written Composition

1. Successful writing is achieved only if all themes are carefully corrected by the teacher.
2. Grades are the most effective way of motivating students to improve their writing.

The Importance of Student Self-Expression in the Instruction of Written Composition

1. Teachers should write all compositions they assign to students.
2. Compositions written in class should never be given letter grades.

The Importance of Linguistic Maturity in the Instruction of Written Composition

1. The experience of composing can and should nurture the pupils' quest for self-realization and their need to relate constructively to their peers.

2. The teacher-pupil conference can and should aid learners in finding their strengths and encourage them in correcting some of their weaknesses.

Validity, Reliability, and Normative Data: Intercorrelations of scores on each of the four scales indicated that scale scores are relatively independent and thus are not measuring a unidimensional attitude. Item-remainder correlations, that correlated response to an item with a score based upon all other items in the scale, further substantiated the convergent and discriminant validity of these scales. Cronbach alpha reliability coefficients for the four scales were computed and confirmed the reliability of the four scales. Normative data are available for twenty-eight teachers who participated in the Puget Sound Writing Program.

Ordering Information: ED 199 717

Related Documents:

Schuessler, Brian F., Anne Ruggles Gere, and Robert D. Abbott. "The Development of Scales Measuring Teacher Attitudes toward Instruction in Written Composition: A Preliminary Investigation." *Research in the Teaching of English* 15 (February 1981): 55–63. (EJ 242 212)

———. "Measuring Teacher Attitudes toward Instruction in Writing." In *New Directions in Composition Research*, edited by Richard Beach and Lillian Bridwell. New York: Guilford Press, 1984.

Category: Teacher Knowledge/Attitudes

Title: Rating Scale for the Assessment of the Speaking Skills of Teachers

Author: Sandra L. Snyder

Age Range: Adult

Description of Instrument:

Purpose: To assess the speaking skills of college students enrolled in teacher education programs.

Date of Construction: 1981

Physical Description: This 100-point scale measures selected components of the speech act judged to be important for effective oral communication by researchers in speech communication and teacher education, and by the author. The scale combines the use of holistic and primary trait scoring currently used in the evaluation of writing. The instrument follows.

Rating Form

Key:
1—poor
2—below average
3—above average
4—very good
✓—each descriptor where performance is acceptable

Time: Start ___ Finish ___
Name: ______________
Topic: ______________

Organization and development 1 2 3 4 (×6) ___

___ Purpose clear
___ Main ideas clear
___ Main ideas consistent with purpose
___ Smooth transitions
___ Logical sequence of ideas
___ Information factual
___ Use of evidence
___ Concluding statement

Adaptation to audience 1 2 3 4 (×6) ___

___ Provides sufficient information
___ Eye contact with audience
___ Relates message to audience
___ Clear explanations

Language usage 1 2 3 4 (×5) ___

___ Use of appropriate vocabulary
___ Use of standard English dialect

___ Enunciation
___ Use of conventional grammar

Ability to motivate audience 1 2 3 4 (×3) ___

___ Personal involvement
___ Speaks expressively
___ Uses variety in presentation
___ Uses visual aids

Delivery 1 2 3 4 (×3) ___

___ Speaks audibly
___ Speaking rate
___ Posture
___ Body movement and gestures

Overall impression 1 2 3 4 (×2) ___

Total score ___

Comments:

Validity, Reliability, and Normative Data: Each component of the scale is derived from a list of behavioral objectives that comprise the traits experts consider essential for effective speaking. The weighting of each component was accomplished by reviewing existing speech rating scales and adapting them based on research in teacher effectiveness. Reliability was demonstrated by two separate tests with two different groups of raters. Both checks were preceded by 2-hour training sessions conducted by the author. Ratings of 5-minute speeches recorded on videotape resulted in a reliability of .93 on both occasions. A paired-judges correlation coefficient was used.

Ordering Information: Full instrument reproduced above.

Related Documents:

Snyder, Sandra L. "An Investigation to Develop and Validate a Rating Scale for the Assessment of the Speaking Competence of Preservice Teachers." Ph.D. diss., Pennsylvania State University, 1981.

Fagan, Edward R., Mary M. Dupuis, and Sandra L. Snyder. "Basic Skills of Prospective Secondary Teachers." *English Record* 32 (Spring 1981): 14–18.

Writing

Category: Writing

Title: Dimensions for Looking at Children's Writings and Drawings in Daily Journals over Time

Author: Amity P. Buxton

Age Range: Ages five through twelve (and above, if drawings are included)

Description of Instrument:

Purpose: To analyze and assess children's spontaneous writing and drawing in daily journals.

Date of Construction: 1976

Physical Description: This instrument provides a framework for an ongoing look at growth in writing. It is designed to aid in the description of writing rather than the evaluation of writing. Nine dimensions cluster around three basic aspects of writing: the thought or meaning (what), the individuality or person of the writer (who), and the form or structure of the writing (how). The instrument follows:

Dimensions for Looking at Children's Writings and Drawings in Daily Journals over Time

(Note: The dimensions are interrelated as the child thinks and writes in his or her journal. The dimensions are isolated here only as analytic tools for a comprehensive view of children's writing and drawing.)

Thought/Meaning—(What)

1. *Theme(s):* Predominant or (over time) recurring subjects or attitudes or feelings or ideas or value judgments . . . motifs . . . aspect of subject consistently discussed . . . interrelationships of recurring subjects, etc.
2. *Organization:* A sequence? undigested? selected? dramatic?—time relationships: overlapping? simultaneous events? —logical thinking? association? . . . generalization? classification? comparison? elaboration? hypotheses? cause-effect? a series? ambiguous? clear? coherent? unrelated? undeveloped?
3. *Vocabulary:* Distinctive choice? precise? varied? limited? predominance of one type? (adverb, adjective, color, action word, ...)

Person—(Who)

4. *Approach or stance of the writer toward the subject:* Personal anecdote? a diary? narrative? imagination? expression? (idea? feeling? value judgment? point of view?)

speaking to an audience? . . . fantasy? "busy work?" . . . information? . . . reflection? record of an event? report? storytelling? pointing out? explaining?

5. *Authorship/uniqueness:* Characteristic constellation of patterns of the child's writing? How can you tell he or she wrote it? . . . "typical" structure? idea? feeling? "style?"
6. *Authenticity/individuality of the writing:* Personal feelings, opinions, perspectives, or ideas which communicate the writer's individuality . . . details which come from personal experience . . . involvement of writer in subject . . . marks of "sincerity." . . .

Form (How)

7. *Interrelationships of writing and drawing:* What does each express about the subject? . . . How do they relate to each other and to total expression/idea/communication? Does writing relate to picture? Which is more detailed? Which took most energy? Does writing have graphic effect? . . . Do they depend upon each other? How do they relate to total page? Which expresses "thought" more clearly? Does writing go beyond picture? What does picture say that words don't?
8. *Language structure:* Complete sentences? simple? compound? complex? varied? Tenses, parts of speech, word order, clauses, subject/verb agreement, etc.; contractions, participles, dummy subject ("it was . . .")? embedding? ambiguous? clear? active or passive voice? awkward? conversational? rhetorical devices related to literary forms: "once upon a time," title, "The End"; conversation; dialogue . . . repetition . . . alliteration; "internalized storytelling"; poetic (word order, phrasing)? devices related to media: journalistic? commercials? comic book dialogue?
9. *Mechanics:* Manual control; underwriting, overwriting, independent handwriting; spelling (and invented spelling); punctuation and signals other than production; unusual use of punctuation (e.g., apostrophes for words other than possessives or contractions); capitalization, upper/lower case confusion.

Validity, Reliability, and Normative Data: Based on research in language, thought, and human development conducted by Piaget, Langer, Cazden, Britton, Vygotsky, and others, the instrument reflects the view that "language is not a body of discrete skills to be learned systematically in a predictable sequence and to be measured periodically against predetermined and standardized expectations." Accordingly, this instrument resulted from a three-year, multi-stage research effort during which data were collected from thirty five through eight-year-olds. Data analysis involved continual reading and rereading of 1,080 samples of writing and drawing in order to identify those features which stood out as significant.

Ordering Information: Full instrument reproduced above.

Related Document:

Buxton, Amity P. "Children's Journals: Further Dimensions of Assessing Language Development." Paper prepared for the Canadian Council of Teachers of English International Conference on Writing, Ottawa, Canada, May 11, 1979.

Category: Writing

Title: Locus of Complexity in Written Language (LCWL)

Authors: Roger L. Cayer
Renee K. Sacks

Age Range: Junior high through adult

Description of Instrument:

Purpose: To analyze the nature and placement of complex syntactic structures in written language as a means of identifying characteristically oral syntactic patterns in writing.

Date of Construction: 1979

Physical Description: The locus of complexity—the placement of modification and elaboration—within a sentence represents an important distinction between oral and written language. Developmentally, inner speech tends to eliminate or drastically abbreviate the subject while stressing the predicate. This is also characteristic of external speech when the subject being discussed is shared. Writing, a mode in which knowledge of the subject shared between the writer and the reader can only be assumed at the risk of causing misunderstanding, tends generally to be more balanced in its syntactic development and organization. The first part of the instrument, a set of measures of syntactic complexity, is drawn from Kellogg Hunt's research in written syntax. It is comprised of the following indexes: (1) mean T-unit length; (2) number, length, and type of subordinate clauses; (3) length and type of phrases—adjective and adverb; and (4) number and type of single-word modifiers. To analyze the locus of complexity in a written sample, first the syntactic factors noted above are identified, and frequency counts are calculated for each as they occur in the subject and predicate portions of each T-unit examined. Second, scores for each individual variable are calculated as a proportion of that variable's occurrence in either the subject or predicate portion of the T-units to the total occurrence of that variable in the corpus. Following is the language elicitation task for the instrument:

> Elaine, a full-time secretary for an insurance company, is twenty-three years old, is married, and has a young son. Her husband, Larry, twenty-five years old, is a construction worker who has been unemployed for the last six months. During this period, he has been taking care of their young son. Encouraged by the hope of a better position and increased pay at the insurance firm, Elaine is thinking of taking a leave of absence from work and returning to college. Her husband Larry opposes this action.

> Sandra, twenty-eight years old, is on the verge of leaving her husband, John, who is thirty-two years old. Both are successful professionals: she is an attorney and he is an accountant. Because of the demands of her job, Sandra wants more freedom away from home; that is, she wants to be able to travel alone, work late hours and go to lunch with some of her clients. John objects. He wants Sandra to reduce her professional commitments and have a child.
>
> As the situations described above suggest, the roles of men and women in our society are changing radically. Increasingly, men are being asked to share in the household chores and in child-rearing so that women can go to work. What are your views and feelings about the newly emerging roles of men and women in our culture today?

Validity, Reliability, and Normative Data: The validity of the measures used to identify and describe syntactic complexity comes from their basis in research by Hunt. Validity of the locus of complexity measures rests on the research by Christensen on the rhetoric of the sentence, Wall's work on predication, and Vygotsky's on inner speech. While no reliability or normative data are reported, the results of several studies utilizing this instrument conducted by the authors demonstrate that the locus of complexity measure significantly distinguishes between the syntactic patterns of oral and written language.

Related Document:

Cayer, R. L., and R. K. Sacks. "Oral and Written Discourse of Basic Writers: Similarities and Differences." *Research in the Teaching of English* 13 (May 1979): 121-28. (EJ 204 587)

Category: Writing

Title: Dichotomous Scale for Evaluating Expository Writing (DSEEW)

Author: Arthur M. Cohen

Age Range: Senior high through postsecondary-adult

Description of Instrument:

Purpose: To evaluate and score expository writing.

Date of Construction: 1973

Physical Description: DSEEW is a dichotomous (yes/no) scale for evaluating expository writing. It was developed to score essays in a curriculum evaluation study at fourteen community colleges in five states. DSEEW has nineteen items, grouped under categories of content, organization, and mechanics. Scorers indicate whether papers do or do not have the feature identified in each item. DSEEW is reproduced below:

	YES	NO		
Content I.	___	___	1.	Ideas themselves are insightful.
	___	___	2.	Ideas are creative or original.
	___	___	3.	Ideas are rational or logical.
	___	___	4.	Ideas are expressed with clarity.
Organization II.	___	___	5.	There is a thesis.
	___	___	6.	Order of thesis idea is followed throughout the essay.
	___	___	7.	Thesis is adequately developed.
	___	___	8.	Every paragraph is relevant to the thesis.
	___	___	9.	Each paragraph has a controlling idea.
	___	___	10.	Each paragraph is developed with relevant and concrete details.
	___	___	11.	The details that are included are well ordered.
Mechanics III.	___	___	12.	There are many misspellings.
	___	___	13.	There are serious punctuation errors.
	___	___	14.	Punctuation errors are excessive.
	___	___	15.	There are errors in use of verbs.
	___	___	16.	There are errors in use of pronouns.
	___	___	17.	There are errors in use of modifiers.
	___	___	18.	There are distracting errors in word usage.
	___	___	19.	The sentences are awkward.

Validity, Reliability, and Normative Data: DSEEW was developed by a group of twenty-one community college English instructors. All twenty-one teachers agreed that the nineteen scales identify the important qualities of expository essays. The special feature of DSEEW is that it asks only whether a quality is present or absent in the writing. Its developers argued that if a quality cannot be judged to be present or absent (hence yes/no or dichotomous scoring), then more than one quality is being assessed in each decision, rendering the scale unreliable.

For each of the nineteen items in DSEEW, rater agreement for fifteen instructors ranged from 53 percent to 100 percent. Internal consistency of scales with their category scores ranged from a correlation of .59 to .73 within the content category, .61 to .78 in organization, and .44 to .59 in mechanics.

Ordering Information: Full instrument reproduced above.

Related Document:

Cohen, A. M. "Assessing College Students' Ability to Write Compositions." *Research in the Teaching of English* 7 (1973): 356–71.

Category: Writing

Title: Evaluating Information in Composition (EIC)

Authors: Committee on Teaching and Its Evaluation in Composition of the Conference on College Composition and Communication

Age Range: Senior high to postsecondary-adult

Description of Instrument:

Purpose: To evaluate high school or college writing programs and writing instruction.

Date of Construction: 1982

Physical Description: EIC includes six different questionnaires that provide information about the foundations of the writing program; the teachers' assumptions, goals, and plans; classroom activities as observed by colleagues; the quality of writing assignments; the quality of teachers' comments on student writing; and the students' perceptions of the quality of instruction.

Validity, Reliability, and Normative Data: EIC was designed by specialists in college writing instruction to provide comprehensive information about many aspects of a writing program. The developers of EIC argue that all widely used questionnaires for collecting students' perceptions of teaching are inappropriate for writing courses.

No reliability or normative data are reported for EIC.

Ordering Information: ED 236 634

Related Document:

CCCC Committee on Teaching and Its Evaluation in Composition. "Evaluating Instruction in Writing: Approaches and Instruments." *College Composition and Communication* 33 (1982): 213-29. (EJ 265 676)

Category: Writing

Title: Writing Apprehension Test

Authors: John A. Daly
Michael Miller

Age Range: Junior high and older (alternate form for elementary)

Description of Instrument:

Purpose: To identify those students who are highly anxious about writing.

Date of Construction: 1975

Physical Description: The twenty-six-item measure is composed of a series of statements about feelings a person has about writing. Respondents indicate the degree to which they agree or disagree with each statement using a Likert-type scale format with five possible responses. With high school or older groups, the test takes about 10 minutes to complete. Test directions and items follow:

> Directions: Below are a series of statements about writing. There are no right or wrong answers to these statements. Please indicate the degree to which each statement applies to you by circling whether you (1) strongly agree, (2) agree, (3) are uncertain, (4) disagree, or (5) strongly disagree with the statement. While some of these statements may seem repetitious, take your time and try to be as honest as possible. Thank you for your cooperation in this matter.
>
> 1. I avoid writing.
> 2. I have no fear of my writing being evaluated.
> 3. I look forward to writing down my ideas.
> 4. I am afraid of writing essays when I know they will be evaluated.
> 5. Taking a composition course is a very frightening experience.
> 6. Handing in a composition makes me feel good.
> 7. My mind seems to go blank when I start to work on a composition.
> 8. Expressing ideas through writing seems to be a waste of time.
> 9. I would enjoy submitting my writing to magazines for evaluation and publication.
> 10. I like to write my ideas down.
> 11. I feel confident in my ability to clearly express my ideas in writing.
> 12. I like to have my friends read what I have written.
> 13. I'm nervous about writing.
> 14. People seem to enjoy what I write.
> 15. I enjoy writing.
> 16. I never seem to be able to clearly write down my ideas.
> 17. Writing is a lot of fun.

18. I expect to do poorly in composition classes even before I enter them.
19. I like seeing my thoughts on paper.
20. Discussing my writing with others is an enjoyable experience.
21. I have a terrible time organizing my ideas in a composition course.
22. When I hand in a composition I know I'm going to do poorly.
23. It's easy for me to write good compositions.
24. I don't think I write as well as most other people.
25. I don't like my compositions to be evaluated.
26. I'm no good at writing.

Validity, Reliability, and Normative Data: The measure has been used in over forty studies. In virtually every investigation it has been found highly reliable (using internal consistency estimates, test-retest correlations, and other procedures). Average internal consistency estimates range from .88 to .95. The validity of the measure has been established in a number of tests. Examples include:

a. The test predicts occupational and academic choices that vary in writing requirements.

b. The test predicts the judged quality of writing. Highly anxious writers encode messages lower in perceived quality than their counterparts.

c. The test predicts enrollment in advanced writing courses.

d. The test differentiates between males and females. Males are slightly more apprehensive than females.

e. The test correlates positively with other, less well-validated measures of writing attitudes, as well as with measures of reading attitudes and speaking attitudes (shyness, stage fright).

f. There is an inverse and significant relationship between writing apprehension, as measured by the test, and self-esteem, both in general and as a writer.

g. The measure is correlated with performance on standardized measures of writing competency (including SAT, ACT, TSWE, ECT, and a number of more local measures).

h. The measure is correlated with quantitative indexes of writing (e.g., number of words, sentences, modifiers, T-units, etc.).

Ordering Information: Full instrument reproduced above.

Related Document:

Daly, John A., and Michael D. Miller. "The Empirical Development of an Instrument to Measure Writing Apprehension." *Research in the Teaching of English* 9 (Winter 1975): 242–49. (EJ 137 933)

Category: Writing

Title: Emig-King Writing Attitude Scale for Students (WASS)

Authors: Janet Emig
Barbara King

Age Range: Junior and senior high

Description of Instrument:

Purpose: To assess students' attitudes toward writing.

Date of Construction: 1979

Physical Description: WASS is a revision of the "Emig Writing Attitude Scale" (Student Version) constructed in 1977 for the New Jersey Writing Project. The revised scale contains forty items. The items in the revised scale represent three categories: preference of writing, perception of writing, and process of writing. Approximately 30 minutes are required to administer the scale that asks students to circle one of five points ranging from "almost always" to "almost never." Sample items:

> Perception of Writing: Good writers spend more time revising than poor writers.
>
> Almost always Often Sometimes Seldom Almost never
>
> Process of Writing: I voluntarily reread and revise what I've written.
>
> Almost always Often Sometimes Seldom Almost never
>
> Preference of Writing: I write letters to my family and friends.
>
> Almost always Often Sometimes Seldom Almost never

Validity, Reliability, and Normative Data: The original scale was used in a 1977–1978 study of the New Jersey Writing Project. Twenty-five teachers and 1,600 students participated in this study. The present scale was constructed after suggested revisions were made by English education graduate students and secondary teachers of English. The suggestions of these experts contributed to the content validity of the scale. Cronbach alpha reliabilities for the subscales of WASS are as follows:

Perception, .589; Process, .726; and Preference, .716.

Ordering Information: ED 236 630

Related Document:

King, Barbara. "Two Models of Analyzing Teacher and Student Attitudes toward Writing: The Emig Attitude Scale and the King Construct Scale." Ph.D. diss., Rutgers University, 1979.

Category: Writing

Title: Analyzing Cohesive Ties (ACT)

Author: Carolyn Hartnett

Age Range: Primary through postsecondary–adult

Description of Instrument:

Purpose: To identify and tabulate the various cohesive ties in writing.

Date of Construction: 1980

Physical Description: ACT is a tabulation sheet for recording all types of cohesive ties in writing. There are columns for recording the cohesive word, the number of the sentence in which it appears, the referent of the cohesive word, and the number of the sentence in which it appears. ACT also provides brief descriptions with examples of the types of cohesion. It includes instructions to tabulators.

Validity, Reliability, and Normative Data: ACT is based on a comprehensive description of the cohesive system in English in Halliday and Hasan's *Cohesion in English* (1977). This book provides a more complete explanation of the types of cohesion than ACT. Hence, an evaluator or researcher considering ACT should read *Cohesion in English* for interpretation.

After 15 hours of training, high school English teachers achieved an interrater reliability of .78 with ACT.

Ordering Information: ED 236 654

Related Documents:

Hartnett, Carolyn G. *Measuring Writing Skills*. Texas City, Tex.: College of the Mainland, 1978. (ED 013 371)

Halliday, M. A. K., and R. Hasan. *Cohesion in English*. London: Longman, 1976.

Category: Writing

Title: Scoring Guides for Children's Writing (SGCW)

Authors: Ann Humes
Bruce Cronnell
Joseph Lawlor
Larry Gentry
Carolyn Fieker

Age Range: Primary and intermediate

Description of Instrument:

Purpose: To evaluate various forms of children's writing: description, narration, exposition, persuasion, personal and business letters, and poems.

Date of Construction: 1980

Physical Description: SGCW is a large and comprehensive battery of scoring guides for children's writing. Like the scoring guides from the National Assessment of Educational Progress reviewed in this section, each scoring guide in SGCW is based on a specific writing task. As a consequence, the content or discourse requirements in students' writing can be central to a comprehensive scoring guide that provides information about what the student did or did not do in the particular writing situation. Each guide also permits scoring of formal aspects of the writing: paragraph indentation, margins, capitalization, punctuation, spelling, and handwriting legibility. Vocabulary in the writing task and directions are carefully controlled for appropriateness in grades 2–6.

Each guide in SGCW includes a writing task and specific directions that emphasize the discourse requirements of the task. There is a discussion of the task for scorers, highlighting the main requirements. There is a scoring key (an actual scoring form) and a scoring guide that specifies criteria for each scale in the scoring key. Scoring is on a three point scale: good, acceptable, unacceptable.

One complete guide is reproduced below. Designed for use in fourth grade, it is based on an expository writing task that requires the student to explain a simple procedure—following a map.

SGCW: Following a Map

Sample Item

Directions: Study the map. Find the library and the museum. Write directions that tell a new person in town how to get from the library to the museum.

Write the directions in an order that is easy to follow.

Do not leave out any important information about which way to go or what to watch for along the way.

map stimulus with landmarks and
with library and museum in boldface
and at opposite ends of the map

Comment: The words *important* and *information* may not be in students' reading vocabularies at this grade level.

Item Description

Using expository-writing skills, the student writes the map directions elicited by the stimulus. The stimulus is a map containing landmarks likely to be familiar to the student. Only one route between start and finish points should be possible in any map stimulus. The specific instructions to students are used (1) to help ensure that students will produce those features important for the task, and (2) to facilitate the construction of an appropriate scoring key that evaluates these features.

Scoring Key

Scoring Criteria	*Good*	*Acceptable*	*Unacceptable*
Content:			
Includes critical information about landmarks.	______	______	______
Includes important information about direction of movement.	______	______	______
Writes directions in a logical order.	______	______	______
Uses precise language.	______	______	______
Limits the paragraph to one main idea.	______	______	______
Form:			
Uses correct grammar and complete sentences.	______	______	______
Capitalizes and punctuates correctly.	______	______	______
Spells correctly.	______	______	______
Writes legibly, with appropriate margins and indentation.	______	______	______

Scoring Guide

Content:

Includes critical information about landmarks [specific critical landmarks are noted here for the scorer's use].

Good: Most critical landmarks are referenced.
Acceptable: Some critical landmarks are referenced.
Unacceptable: Few or no landmarks are referenced.

Includes important information about direction of movement [specific information about direction is noted for the scorer's use].

Good: Enough directional information is given to provide the reader with all necessary knowledge about movement.
Acceptable: Some directional information is given.
Unacceptable: Little or no directional information is given.

Writes directions in a logical order.

Good: All information is given in sequential order.
Acceptable: Most information is given in sequential order.
Unacceptable: Little or no information is given in sequential order.

Uses precise language.

Good: Precise words are used to give directions, e.g., "north" or "left."
Acceptable: General words are used to give directions, e.g., "down the street."
Unacceptable: Language used does not convey a sense of direction.

Limits the paragraph to one main idea.

Good: All or most sentences pertain to following the map.
Acceptable: Many sentences pertain to following the map.
Unacceptable: Few or no sentences pertain to following the map.

Form:

Uses correct grammar and complete sentences.

Good: The composition has few or no grammatical errors and all sentences are complete.
Acceptable: The composition has some grammatical errors and most sentences are complete.
Unacceptable: The composition has many grammatical errors and many sentences are fragments and/or run-ons.

Capitalizes and punctuates correctly.

Good: The composition has few or no errors.
Acceptable: The composition has some errors.
Unacceptable: The composition has many errors.

Spells correctly.

Good: Most words are correctly spelled.
Acceptable: Several different words are misspelled.
Unacceptable: Many different words are misspelled.

Comment: The ratings refer to "different words" because many instances of misspelling the same word should be evaluated as one misspelled word.

Writes legibly, with appropriate margins and indentation.

Good: The composition has few or no exceptions to the criterion.
Acceptable: The composition has some exceptions to the criterion.
Unacceptable: The composition has many exceptions to the criterion.

Validity, Reliability, and Normative Data: No reliability or normative data are available in the document in which SGCW appears. With careful training of scorers, teachers or researchers could expect high scorer agreement. SGCW has strong construct validity because it assesses a wide range of features in writing, with a balance between form and content. It is based on current discourse theory as well as actual assignments in current language arts texts used in elementary schools. Since the various scoring guides in SGCW assess nine different kinds of writing appropriate for grades 2-6, evaluators can insure content validity by selecting guides to match kinds of writing taught at the level to be assessed.

Ordering Information: ED 192 371, ED 192 372, ED 192 373,
ED 192 374, ED 192 375, ED 192 376

Category: Writing

Title: Cohesion Scoring Guides (CSG)

Authors: National Assessment of Educational Progress Consultants and Staff

Age Range: Primary through postsecondary-adult

Description of Instrument:

Purpose: To score student writing for cohesion and coherence.

Date of Construction: 1977

Physical Description: CSGs are a set of writing tasks and scoring guides that permit an analysis and scoring of essays on the basis of their cohesion and coherence. Scoring is on a 5-point scale. Three of these paired CSGs were developed for the 1978–79 National Assessment of Writing, one guide for each age level, nine, thirteen, and seventeen. CSGs focus only on the cohesion of the essay, not on rhetorical traits or on syntax, usage, or mechanics. One of the CSGs (for age nine) is reproduced here. The other two are available in documents referenced below.

CSG: Writing Task ("Fireflies")

> (The writing task is based on a picture of a girl catching fireflies. She has a jar with fireflies in one hand, and with the other hand she is reaching for a firefly. The picture is reproduced on page 83 of the document listed under "Ordering Information" below.)
>
> > Here is a picture of a girl who is having fun in the summer. Look at the picture for a while. What do you think she is doing? What do you think she might do next?
> >
> > Write a story that tells what the picture is about.

CSG: Scoring Guide ("Fireflies")

> In scoring papers for cohesion, scorers need to be attentive not only to the incidence of cohesive ties but also to their successful ordering. Underlying and further strengthening these ties is syntactic repetition, both within and across sentences. The following example achieves cohesion by lexical cohesion, conjunction, reference, and substitution, and yet these various kinds of cohesion are both emphasized and related among themselves by numerous incidents of syntactic repetition:
>
> > There is a girl who is catching fireflies. She is putting some into a jar. When she is finished, she will take them into a dark room and watch them glow. After that she will let them go so that they could lay eggs and there will be more fireflies for next year. Then she can catch them again year after year.

When both the incidence and ordering of cohesive ties pattern the entire piece of writing, the writer has created what we ordinarily call coherence.

Scoring Guide Categories:

1 = *Little or no evidence of cohesion.* Basically, clauses and sentences are not connected beyond pairings.

2 = *Attempts at cohesion.* There is evidence of gathering details but little or no evidence that these details are meaningfully ordered. In other words, very little seems lost if the details are rearranged.

3 = *Cohesion.* Details are both gathered and ordered. Cohesion is achieved in the ways illustrated briefly in the definition above. Cohesion does not necessarily lead to coherence, to the successful binding of parts so that the sense of the whole discourse is greater than the sense of its parts. In pieces of writing that are cohesive rather than coherent, there are large sections of details which cohere but these sections stand apart as sections.

4 = *Coherence.* While there may be a sense of sections within the piece of writing, the sheer number and variety of cohesion strategies bind the details and sections into a wholeness. This sense of wholeness can be achieved by a saturation of syntactic repetition throughout the piece (see description above) and/or by closure which retrospectively orders the entire piece and/or by general statements which organize the whole piece.

7 = Illegible, illiterate.

8 = Misunderstands the task, writes on another topic.

9 = "I don't know."

Note: Scorers should not take mechanics or transcription errors into consideration. Also, the scorers should judge only the *interrelatedness* of the ideas, *not* the quality of those ideas.

Examples of Cohesive Ties:

In general, "cohesion" refers to the ways clauses and sentences are related to each other and can be thought of as the gathering and ordering of related ideas. If the parts of a discourse cohere, they "stick" or are "bound" together. Cohesion is achieved by ties of considerable variety. And these ties can be both semantic and structural. Additional examples of specific kinds of cohesive ties are identified by Halliday and Hasan in *Cohesion in English* (1976).

Lexical

The girl has a jar to put *bugs* in. The *bugs* are called fireflies.

Conjunction

Additive—
The girl is catching lightning bugs. She is *also* catching butterflies.

Adversative—
I wanted to help the little girl catch fireflies, *but* I couldn't find her.

Causal—
This little girl is trying to catch fireflies *so* she can take them to school.

Temporal—
She is catching lightning bugs and putting them in a jar. *Next* she will show them to her mother. *Later* she might let them go.

Reference

Personal—
There once was a girl. *She* liked to catch bugs.

Demonstrative—
She is collecting bugs. *This* collection is for her science class.

Comparative—
I wish I had some bubbles *like* hers.

Substitution

Nominal—
The lightning bugs are out and the little girl wants to catch *some*.

Clausal (use of *so* and *not*)—
The little girl knows they are fireflies because her mother said *so*.

Ellipsis

Nominal—
The girl's mother told her to let the bugs go but she wouldn't [ ].

Verbal—
She had to go to her room and couldn't come out until her mother said she could [ ].

Clausal—
She is catching either lightning bugs or butterflies but I don't know which [ ].

Note: While helping plan the 1978–79 writing assessment, National Assessment consultants expressed the opinion that coherence and cohesion deserved special consideration and that a more thorough method of describing information about coherence was needed. In consequence, this cohesion scoring guide was developed and used with this exercise to replace the paragraph coherence guidelines developed in 1973–74.

Validity, Reliability, and Normative Data: CSGs are based on current linguistic research into the nature of cohesion in English (Halliday and Hasan 1976).

After careful training of raters, the scoring contractor for National Assessment of Educational Progress was able to sustain rater agreement above 90 percent with CSGs.

Since CSGs have been used in three National Writing Assessments, a large amount of normative data is available from ages nine, thirteen, and seventeen. For example, in 1979 on the "Fireflies" task, 29 percent of nine-year-olds' essays were rated inadequate (score point 1 on the previously reproduced guide), 46 percent made attempts at cohesion (score 2), 21 percent were cohesive (score 3), and only 1 percent (score 4) displayed both cohesion and coherence.

Ordering Information: ED 205 583

Related Documents:

Mullis, I. V. S., and J. Mellon. *Guidelines for Describing Three Aspects of Writing: Syntax, Cohesion, and Mechanics*. Denver: National Assessment of Educational Progress, 1980. (ED 205 572)

Halliday, M. A. K., and R. Hasan. *Cohesion in English*. London: Longman, 1976.

Category: Writing

Title: Guidelines for Categorizing Mechanics and Grammatical Errors (GCMGE)

Authors: National Assessment of Educational Progress Consultants and Staff

Age Range: Primary through postsecondary-adult

Description of Instrument:

Purpose: To describe and classify the full range of mechanical and grammatical errors in writing.

Date of Construction: 1977

Physical Description: GCMGE is a guide for comprehensive analysis and classification of errors in syntax, usage, mechanics, and spelling. Sentences are first classified as to type of construction, then sentence errors (fused, run-on, comma splice, fragment) are noted. Within the sentence, GCMGE classifies a wide range of usage problems: agreement, reference, tense shift, and ambiguous modification. It also includes classifications for spelling, word choice, capitalization, and all forms of punctuation.

Validity, Reliability, and Normative Data: As a descriptive instrument, GCMGE provides a system for noting virtually any sentence-level error that might appear in students' essays. It is valid in that it is based on current linguistic research and is derived empirically from analyses of a great many essays written by nine-, thirteen-, and seventeen-year-olds. Anyone who has ever attempted systematic error analysis of a writing sample will recognize the value of an empirically based, field-tested guide that fully anticipates all the unpredictabilities of error analysis.

Though GCMGE in its present form was not devised until 1977, just prior to the third National Writing Assessment, it was used to describe writing samples from nine-, thirteen-, and seventeen-year-olds in three National Writing Assessments. Hence, a large amount of normative data is available. For example, a typical descriptive essay written by a seventeen-year-old in 1979 had .3 sentence fragments, .6 run-on sentences, 4.1 misspelled words, and 3.2 punctuation errors, and nearly all of them had omissions of commas.

Ordering Information: ED 205 572

Category: Writing

Title: Guides for Evaluating Expressive Writing (GEEW)

Authors: National Assessment of Educational Progress Consultants and Staff

Age Range: Primary through postsecondary-adult

Description of Instrument:

Purpose: To describe and evaluate the primary rhetorical traits of expressive and expressive/narrative writing.

Date of Construction: 1977

Physical Description: GEEWs are a set of expressive writing tasks for ages nine, thirteen, and seventeen, with a scoring guide for each task. Paired writing tasks and scoring guides are designed to assess writing performance in only the *expressive* rhetorical mode or discourse type. The scoring guide focuses only on the central rhetorical criterion or "primary trait" of expressive writing—the ability of the student to express values and feelings and to elaborate the expression of feeling with anecdote or illustration. The guides do not include criteria for syntax, usage, or mechanics.

One writing task and its scoring guide (designed for students aged thirteen) are reproduced here. Other GEEWs are available in documents referenced below.

GEEW: Writing Task ("Loss")

> Everybody knows or can imagine what it is like to lose something or someone of special importance. Valuable things may be lost or broken, close friends or relatives may die or move away, favorite pets may be lost or killed.
>
> Think of some loss you have experienced. Tell what you especially remember about what you lost, and how it feels to experience such a loss.

GEEW: Scoring Guide

> *Rhetorical Mode:* Expressive
>
> *Primary Trait:* Expression and substantiation of value and feeling through recollection and inventive elaboration.
>
> *Rationale of Primary Trait:* This exercise is oriented to writing about the experience of loss; in particular, the kind of loss which arouses intense feeling. The directive for the exercise requires respondents to write about the loss in two interrelated ways. First, respondents are asked to "tell what you especially remember about what you lost." In this way, they are led to express and to substantiate the "special importance" of the lost

object, pet, or person. Respondents are next asked to tell "how it feels to experience such a loss." In other words, they have to translate feelings into tangible terms. The directive as a whole requires respondents to use writing as a means of defining the nature of a personal loss—by defining the value of what was lost and by defining the felt experience of losing that object, pet, or person. In both instances, the definition is expressed and substantiated through recollection.

General Scoring Rationale: In rating this exercise, readers should look for evidence that writing is being used to express and substantiate the nature of a particular loss—with respect to *both* the importance of what has been lost *and* the feeling about the loss. The first may be done through (1) connotative or value-laden description of the object, pet, or person; (2) description or narration of shared activities or past events involving the object, pet, or person; (3) metaphoric statements about the relationship between the respondent and what has been lost. The feeling may be established by (1) descriptions of mental, emotional, or physiological reactions to the loss; (2) descriptions of physical reactions to the loss, such as looking for the object, burying the pet, or visiting the grave of a person; (3) metaphoric statements which define or seek to define the feeling by using comparisons. In looking for evidence that both *value* and *feeling* have been expressed and substantiated, readers should not be misled or distracted by pure reporting of events leading up to or circumstances concerning the time and place of the loss. It is inevitable that respondents will include some facts, but readers should recognize that merely factual reports or sections of a response given over to factual reporting are not evidence of a particular value or a particular feeling associated with the loss.

Readers should also be aware that assertions of value, feeling, or reaction—"It was of great value to me," "It was important to me," "I was sad," "I felt bad," "I cried"—are too vague and generalized in and of themselves to be regarded as evidence of substantiation. Readers should look for specific and detailed evidence in the responses that writing is being used to express and substantiate the emotional process that loss involves. Something of value which once existed and produced feelings of pleasure or satisfaction no longer exists. (For responses that consider more than one loss, readers should choose the section of the paper that would receive the highest classification.)

Scoring Guide Categories:

0 = No response.

1 = *Little or no expression of value and feeling.* These responses show no or only vague evidence of using writing to express and substantiate value and feeling through recollection and elaboration of details concerning a particular loss.

a = Some "1" responses offer factual reports of varying lengths but include no or only vague assertions of feeling.

b = Some "1" papers list a series of losses. Some parts of the series identify the losses and nothing more; others may offer vague assertions of feeling.

c = Other "1" responses are just too sparse to provide any substantiation of feeling and/or value. "I was sad when my favorite grandfather died." "My cute puppy was run over by a car. I cried." "I was sad and depressed when my dog died." "My aunt was nice. She came over everyday."

2 = *Moderate expression of value and feeling.* These responses show some evidence of using writing to express and substantiate value and feeling through recollection and inventive elaboration of details concerning a particular loss.

a = Some "2" papers offer details (2–3) to establish and substantiate feeling about the loss but do little or nothing to substantiate the value of what has been lost.

b = Some "2" papers offer details (2–3) to substantiate the value of what has been lost but do little or nothing to substantiate a feeling about the loss. Some of these papers may even substantiate the value of what has been lost at considerable length, but any feeling concerning loss is only vaguely present.

c = Some of these responses substantiate both value and feeling, but the details are few (1 or 2 for each dimension) and relatively generalized.

3 = *Expression of value and feeling.* These papers use writing to express and substantiate value and feeling through recollection and inventive elaboration of details concerning a particular loss.

a = Some "3" papers offer extensive substantiation of feeling, yet they do little or nothing to substantiate the particular value of what has been lost. Still, the feelings expressed in the responses imply the value of the loss.

b = Some papers which substantiate the value of what has been lost at considerable length may also be classified "3" if feeling is implied. The reader should have a real sense of closeness or loss.

c = Some "3" responses substantiate both value and feeling (2 or 3 details for each dimension), but the development is still somewhat uneven or the details tend to be generalized.

4 = *Developed elaboration of expression of value and feeling.* These papers express and substantiate value and feeling at length, and they do so through details that are sufficiently specific and vivid to establish the precise quality of what has been lost and the nature of the feelings experienced about the loss. These papers tend to be well organized and develop the experiential quality of the loss.

7 = Illegible, illiterate.

8 = Misunderstands the task, writes on another topic.

9 = "I don't know."

Validity, Reliability, and Normative Data: GEEWs have strong construct validity because they are based on an analysis of the primary rhetorical requirements of expressive writing. They are also based on current discourse theory (Lloyd-Jones 1977). GEEWs would have content validity as a criterion measure in any writing program concerned with the development of students' expressive writing ability.

After careful training of raters, the scoring contractor for National Assessment of Educational Progress was able to sustain rater agreement above 90 percent with GEEWs.

Since GEEWs were used in the second and third National Writing Assessments to score papers from a stratified random sample of nine-, thirteen-, and seventeen-year-olds, a large amount of normative data is available. For example, in 1979 on the "Loss" writing task reproduced above, 40 percent of the thirteen-year-olds' essays were scored as displaying little valuing and feeling (score point 1 on the guide reproduced above), 39 percent some valuing and feeling, 18 percent clear valuing and feeling, and just over 1 percent elaborated valuing and feeling.

Ordering Information: ED 205 583

Related Documents:

Lloyd-Jones, R. "Primary Trait Scoring." In *Evaluating Writing: Describing, Measuring, Judging,* edited by C. R. Cooper and L. Odell. Urbana, Ill.: National Council of Teachers of English, 1977.

Mullis, I. V. S. *Using the Primary Trait System for Evaluating Writing.* Denver: National Assessment of Educational Progress, 1980.

National Assessment of Educational Progress. *Writing Achievement, 1969–1979: Results from the Third National Writing Assessment* (Vol. 1: Seventeen-year-olds; Vol. 2: Thirteen-year-olds; Vol. 3: Nine-year-olds). Denver: National Assessment of Educational Progress, 1980.

Category: Writing

Title: Guides for Evaluating Persuasive Writing (GEPW)

Authors: National Assessment of Educational Progress Consultants and Staff

Age Range: Primary through postsecondary-adult

Description of Instrument:

Purpose: To describe and evaluate the primary rhetorical traits of persuasive writing.

Date of Construction: 1977

Physical Description: GEPWs are a set of persuasive writing tasks for ages nine, thirteen, and seventeen, with a scoring guide for each task. Paired writing tasks and scoring guides are designed to assess writing performance in only the ***persuasive*** rhetorical mode or discourse type. The scoring guide focuses only on the central rhetorical criterion or "primary trait" of persuasive writing—the ability to develop an argument supporting a point of view on a controversial issue. The guide does not include criteria for syntax, usage, or mechanics. One writing task and its scoring guide (designed for students aged seventeen) are reproduced here. Other GEPWs are available in documents referenced below.

GEPW: Writing Task ("Recreation Center")

Some high school students have proposed converting an old house into a recreation center where young people might drop in evenings for talk and relaxation. Some local residents oppose the plan on the grounds that the center would depress property values in the neighborhood and attract undesirable types. A public hearing has been called. Write a brief speech that you would make supporting or opposing the plan. Remember to take only ONE point of view. Organize your arguments carefully and be as convincing as possible.

GEPW: Scoring Guide

Rhetorical Mode: Persuasive—Social/Community

Primary Trait: Persuasion through invention and elaboration of arguments appropriate to specified issues and limited to an audience with a mixed bias.

Rationale of Primary Trait: This task represents controversial situations that prevail in any civilized society—situations which are resolved by a deliberative response. The directive to "be as convincing as possible" indicates the persuasive orientation of the task. It requires that respondents develop and support arguments appropriate to their position.

General Scoring Rationale: Support may consist of evidence and/or appeals to general truths, to experience, or to social and economic values. The support must be consistent with the position and should be of at least moderate length to demonstrate competence (scale point "3"). Excellence is achieved by demonstrating a capacity not only to invert and support arguments but also by addressing both sides of a controversial issue. Thus, the most successful respondents will be able to support their cases on their own merits as well as answer or refute at moderate length the causes of the opposition.

Scoring Guide Categories:

0 = No response.

1 = *Do not define and defend a point of view.* Some of these papers have not explicitly or implicitly taken a position. Others may contain a thesis statement or clearly imply a position but do not give several supporting reasons to develop their arguments. Some typical score point "1" papers present:

a = Attitudes and opinions about related social issues without a clear statement of position—these include free-floating, uncontrolled statements of opinion showing no concern for taking a stand and supporting it.

b = Position statements but no related support—often these papers merely reiterate their stand in various forms.

c = Position statements preceded or followed by elaborate introductions.

d = Position statements followed by arguments and appeals not connected to the crucial issues.

e = Position statements followed by one or two undeveloped reasons.

f = Position statements, but the paper goes off tangentially into another realm (clarifying terms, personal gripes, etc.).

2 = *Define a point of view and offer minimal defense.* These papers explicitly state or strongly imply a position and give one or more clusters of arguments or appeals. (A cluster is a reason asserted with no more than one or two bits of evidence or related appeal.) Score point "2" papers usually consist of a chain of briefly developed appeals in support of a position or answering the opposition. They do not develop a line of argument or link the clusters to each other. (The underlying assumption is that the lines of arguments, reasons, or appeals are appropriate to the issue.)

3 = *Define and defend a point of view.* These papers clearly state or imply a position and present at least one substantially developed line of argument or two moderately developed lines of argument relevant to the issues at hand. More evidence to support the position is presented than in "2" papers.

4 = *Systematically define and defend a point of view.* These papers present at least two moderately developed lines of argument, one which supports the position and one which answers the possible arguments raised by the opposition. The lines of argument usually will be linked as well as carefully organized. Other "4" papers may contain a moderate statement of support with a brief address answering each of the major opposition positions.

7 = Illegible, illiterate.

8 = Misunderstands the task, writes on another topic.

9 = "I don't know."

Validity, Reliability, and Normative Data: GEPWs have strong construct validity because they are based on an analysis of the primary rhetorical requirements of persuasive writing. They are also based on current discourse theory (Lloyd-Jones 1977). GEPWs would have content validity as a criterion measure in any writing program concerned with the development of students' persuasive writing ability.

After careful training of raters, the scoring contractor for National Assessment of Educational Progress was able to sustain rater agreement above 90 percent with GEPWs.

Since GEPWs were used in the second and third National Writing Assessments to score papers from a stratified random sample of nine-, thirteen-, and seventeen-year-olds, a large amount of normative data is available. For example, in 1979 on the "Recreation Center" writing task reproduced above, 25 percent of the seventeen-year-olds' essays were scored as not persuasive (score point 1 on the guide reproduced above), 56 percent as minimally persuasive (score 2), 15 percent as persuasive (score 3), and less than 1 percent as fully persuasive.

Ordering Information: ED 205 583

Related Documents:

Lloyd-Jones, R. "Primary Trait Scoring." In *Evaluating Writing: Describing, Measuring, Judging,* edited by C. R. Cooper and L. Odell. Urbana, Ill.: National Council of Teachers of English, 1977.

Mullis, I. V. S. *Using the Primary Trait System for Evaluating Writing.* Denver: National Assessment of Educational Progress, 1980.

National Assessment of Educational Progress. *Writing Achievement, 1969–1979: Results from the Third National Writing Assessment* (Vol. 1: Seventeen-year-olds; Vol. 2: Thirteen-year-olds; Vol. 3: Nine-year-olds). Denver: National Assessment of Educational Progress, 1980.

Category: Writing

Title: Sentence Combining Scoring Guides (SCSG)

Authors: National Assessment of Educational Progress Consultants and Staff

Age Range: Primary through postsecondary–adult

Description of Instrument:

Purpose: To describe and score students' responses to sentence-combining test items.

Date of Construction: 1977

Physical Description: SCSGs present students with sentence-combining exercises and then permit a comprehensive description and scoring of the students' combined sentences. Scoring takes account of combinings into a single T-unit (independent clause and all related modifying clauses or phrases), inclusion of appropriate lexical content, and correctness of syntax. In addition, SCSGs allow a classification of the combining strategies students used.

Altogether, fourteen separate combining exercises were developed for the 1978–79 National Assessment of Writing. They are available in the document listed under "Ordering Information" below.

Reproduced here are the standard instructions for all the exercises and three exercises administered to students aged thirteen and seventeen.

SCSG: Instructions and Three Exercises

Below are some sets of short sentences. Each set can be improved by combining the given sentences into *one* sentence that says the same thing. For example, if the sentences were:

A cat chased the ball.
The cat was big.
It was gray.

You could write:

a big gray cat chased
the ball.

After you hear each set read aloud, read the sentences silently to yourself and figure out a way to combine them into one sentence. Be sure your sentence has the same meaning as the sentences in the given set. Then write your sentence on the lines. Now here is the first set of sentences to be combined:

A. The boys drank the lemonade.
The boys were barefoot.
The lemonade was cold.

____

____

____

B. The pebbles marked the path to a kingdom.
The pebbles were shiny.
The pebbles were yellow.
The pebbles were gleaming like cats' eyes.
The kingdom was magic.
The kingdom was underground.
The kingdom was ruled by a wizard.

____

____

____

____

C. The hikers tramped along the path.
The path was steep.
It was narrow.
It was rocky.
It curved upward toward the mountain top.
The mountain top appeared ahead through the clouds.

____

____

____

____

Validity, Reliability, and Normative Data: SCSGs are intended to test students' competence in applying the range of sentence-combining options in English. Construct validity of SCSGs are supported by the fact that older students did better than younger students on the same exercises in the 1978-79 National Writing Assessment. SCSGs would have content validity for a school writing program which emphasized syntactic fluency and flexibility. It is not a test of composing or writing but only of certain important syntactic operations.

Normative data are available only on individual exercises (Mellon 1981). For example, on the three exercises illustrated above, the combined national results were as follows:

Exercise	*Age*	*Percent Correctly Combined into 1 T-unit*	*Percent Nonratable*
Lemonade	9	34	30
	13	77	7
	17	91	2
Pebbles	9	3	33
	13	21	8
	17	48	2
Hikers	13	23	7
	17	44	2

No reliability data are available. However, since SCSGs are based on close linguistic analysis and description, one can assume that trained raters could sustain a high percentage of agreement on individual sentence-combining exercises.

Ordering Information: ED 205 583

Related Document:

Mellon, John. *Sentence-Combining Skills: Results of the Sentence-Combining Exercises in the 1978–79 National Writing Assessment.* Denver: National Assessment of Educational Progress, 1981.

Category: Writing

Title: Scale for Evaluating Expository Writing (SEEW)

Author: Edys Quellmalz

Age Range: Intermediate through postsecondary-adult

Description of Instrument:

Purpose: To evaluate and score expository writing.

Date of Construction: 1981

Physical Description: SEEW provides separate 6-point rating scales for general impressions of the quality of an essay, the general competence, the coherence, the paragraph organization, the support for main ideas, and the mechanics. The mechanics scale includes a brief guide for identifying certain errors in sentence construction, usage, spelling, punctuation, and capitalization. Designed for scoring essays from competency or proficiency tests, SEEW defines score points 4–6 in terms of mastery or competence, score points 1–3 in terms of non-mastery.

As an example of one scale in SEEW, the complete scale for coherence is presented here:

SEEW: Essay Coherence

This subscale focuses on the flow of ideas throughout the entire paper and *between* paragraphs. The emphasis is on *vertical* relationships of ideas throughout the essay.

Essay coherence: The paper has a main idea (stated or clearly implied) that makes a point about the subject and is at a greater level of generality than the other points within the paper. Subtopics are logically related to the main idea and to each other.

Master

6 =
- The subject is identified
- The main idea is stated or implied in opening and/or closing statement
- Opening and closing statements must match or logically relate to the text and to each other
- The topic is limited through reference to key points or lines of reasoning
- The essay plan is logical
- The essay plan is clearly signaled by transitions
- The essay plan is consistently maintained (no digression or extraneous material)

5 =
- The subject is identified
- The main idea is stated or implied in opening and/or closing statements

- Opening and closing statements relate to or follow from the text and from each other
- The topic is partly limited by indicating number and type of key points
- Plan is logical
- Plan is signaled by appropriate transitions
- There may be digression or an elaboration

4 =
- The subject is identified
- The main idea is identified or implied
- There may or may not be an attempt to limit the topic, give directions to subsequent reasoning
- There may be a few minor digressions from the plan, but no major digressions
- Subtopics can be reshuffled

Non-Master

3 =
- Subject is clear
- Main point may *not be very clear.* There may be a major digression or several minor digressions
- A plan is attempted that may need to be inferred

2 =
- Subject is clear
- Main idea is not very clear and/or there may be more than one
- The plan is attempted, but not consistently or not completely carried out
- There are many digressions

1 =
- Subject is unclear
- Main idea is absent or very unclear
- No plan is attempted or followed

Validity, Reliability, and Normative Data: SEEW is a composite of several current approaches to general impression and analytic scoring. Choices could be made among its six scales to reflect instructional emphases in a school program. Reliability data come from six raters scoring over 1,000 essays. After a careful training session, coefficients of reliability ranged from .63 to .77 on the subscales of each scale in SEEW.

Ordering Information: ED 236 670

Related Document:

Cooper, Charles R. "Holistic Evaluation of Writing." In *Evaluating Writing: Describing, Measuring, Judging,* edited by C. R. Cooper and L. Odell. Urbana, Ill.: National Council of Teachers of English, 1977. (ED 143 020)

Category: Writing

Title: Scale for Evaluating Narrative Writing (SENW)

Author: Edys Quellmalz, Center for the Study of Evaluation, University of California, Los Angeles

Age Range: Elementary through postsecondary-adult

Description of Instrument:

Purpose: To evaluate and score narrative writing.

Date of Construction: 1978–1981

Physical Description: SENW provides separate 6-point rating scales for the general competence of an essay and the levels of development of focus and organization, support, and grammar/mechanics. The mechanics scale includes a brief guide for identifying certain errors in sentence construction, usage, spelling, punctuation, and capitalization. Designed as a criterion-referenced scale to describe levels of writing skill development for *basic* essay elements, SENW defines score points 4–6 in terms of acceptable levels of development and score points 1–3 as below competency.

As an example of one scale in SENW the complete scale for support is presented here:

SENW: Support

> This element focuses on the quality (specificity and amount) of the support provided for the essay theme both within each paragraph and throughout the essay.
>
> Supporting details should be at a greater level of specificity and depth than the generalizations they are intended to develop. Events, descriptions, and characters should be fleshed out through the use of specific details, such as examples, facts, anecdotes, or descriptions. These details should provide the reader with an image of the appearance, feelings, thoughts, actions, or mood of the events taking place in the narrative.
>
> Master
>
> 6 =
> - All events, characters, and/or descriptions are well developed by specific and clear supporting details, such as examples, descriptions, anecdotes, or facts.
> - Supporting statements or details provide in-depth descriptions or statements about appearance, feelings, thoughts, actions, and/or mood.
> - Supporting details are more specific than the general ideas, events, or characters they describe.

5 =

- Most events, descriptions, and/or characters are developed through effective use of details, but the use of detail may be slightly uneven. For example, a character's physical appearance may be more fully described than his or her thoughts, emotions, or actions.
- Supporting statements or details provide in-depth descriptions about appearance, feeling, thoughts, actions, and/or mood, but there may be one or two instances in which details lack depth.
- Supporting statements or details clearly are more specific than the general statements they describe.

4 =

- Most events, descriptions, and/or characters are developed through use of details, but the use of details may be slightly uneven.
- Some of the supporting statements or details may not provide sufficient in-depth descriptions about appearance, feeling, thoughts, actions, and/or mood.
- Supporting statements or details clearly are more specific than the general ideas, events, or characters described.

Non-Master:

3 =

- The use of detail is uneven. Several statements or descriptions are not developed through the use of detail.
- Some of the supporting statements are not sufficiently specific and do not provide sufficient in-depth descriptions about the appearance, thoughts, actions, or mood of characters, events, or setting.

2 =

- Many supporting details or statements lack specificity and depth.
- The details are not smoothly integrated in the composition.

1 =

- Supporting statements are vague or confusing references to the events, characters, or ideas they describe.
- There is little or no evidence of supporting details.

Validity, Reliability, and Normative Data: SENW is a composite of several current approaches to holistic and analytic scoring. School systems could make choices among or changes in the six scales to reflect instructional emphases or assessment philosophy. Reliability data come from controlled research and actual competency assessments in which several hundred raters have scored over 10,000 essays. After a careful training session, coefficients of reliability ranged from .89 to .91 on the subscales. Rating time per essay averages 2.3 minutes for paragraph-length compositions, and 3.5 minutes for multi-paragraph essays.

Ordering Information: ED 236 653

Related Documents:

Cooper, C. R. "Holistic Evaluation of Writing." In *Evaluating Writing: Describing, Measuring, Judging,* edited by C. R. Cooper and L. Odell. Urbana, Ill.: National Council of Teachers of English, 1977. (ED 143 020)

Mullis, I. V. S. "Scoring Direct Writing Assessments: What Are the Alternatives?" *Educational Measurement* 3 (1984).

Category: Writing

Title: Scoring Writing with an Informative Aim

Author: Nancyanne Rabianski

Age Range: High school, entrance-level college students

Description of Instrument:

Purpose: To evaluate the degree to which an essay is successful in giving information.

Date of Construction: 1977

Physical Description: This instrument is an analytic/holistic scoring guide usable for evaluating any writing that contains generalizations supported by elaboration. The following criteria are applied: (1) Relevance—any statement that gives information or elaborates on information already given receives credit; (2) Nonrepetition—statements are not accorded credit more than once; (3) Variety—a variety of statements should support the general thesis; (4) Depth—statements should be elaborated by giving a number of supporting examples or a lengthy narration. The holistic judgment is determined first and is based on the raters' initial impressions of the paper. The analytic judgment, which may or may not agree with the holistic rating, is determined by analysis of the degree to which each criterion is met. Each generalization is also scored with one of five numerical values according to the quantity of supportive elaboration employed:

a. Unelaborated Generalization—This is the basic type of information imparted. It is a statement of what is basically believed to be true. No examples, specific generalizations, or narratives are given as explanations of the generalization.

b. Some Elaboration—The writer gives one or two examples or specific generalizations. For example, as a specific generalization a writer might state: "Money is important to my dad. He likes luxuries." As a specific example, a writer might state: "He bought a $670 television."

c. Much Elaboration—The writer gives three or four examples or specific generalizations: "Money is important to my dad. He likes luxuries. He bought a $670 television. He has three CB radios that he bought last summer."

d. Very Much Elaboration—The writer gives five or six brief examples, five or six specific generalizations, or a narrative of five or six clauses. In the case of the narrative, the length of five clauses usually ensures that it is truly a narrative rather than simply an allusion to an incident.

e. Outstanding Elaboration—Seven or more examples, specific generalizations, or narrative clauses.

Validity, Reliability, and Normative Data: Guiding the development of the instrument were the theories of discourse of Kinneavy and Britton. The rating guides are based on an examination of 372 essays produced by 93 tenth graders as part of a doctoral study. Each student wrote four essays that were scored by four raters. The raters were experienced high school teachers, each of whom had received twenty hours of training. Interrater reliability for the total of all topics scored analytically was .885. Normative data are also available from the rating procedure's later use for the analysis of freshman writing at SUNY, Buffalo.

Ordering Information: ED 236 638

Related Documents:

Rabianski, Nancyanne. "An Exploratory Study of Individual Differences in the Use of Free-Writing and the Tagmemic Heuristic Procedure." Ph.D. diss., State University of New York at Buffalo, 1977.

State University of New York, Academic Affairs Office. *Writing Abilities of Regularly-Admitted Freshmen at SUNY/Buffalo.* Buffalo, N.Y.: State University of New York, Fall 1979.

Category: Writing

Title: Writing Center Tutorial Record Form (WCTRF)

Authors: Tom Reigstad
Ann Matsuhashi
Nina Luban

Age Range: Senior high and postsecondary-adult

Description of Instrument:

Purpose: To describe and assess the activities engaged in between a writing tutor and a student in a one-to-one writing conference.

Date of Construction: 1980

Physical Description: WCTRF is two forms. One form, which the student fills out on his or her first visit to a writing center, provides information about the student's major, native language, source of the writing task the student needs help with, and the student's perception of the problem. The second form the tutor fills out immediately after the conference, and it records specific information about the activities during the conference. This information can be easily entered onto computer score sheets, and after several such record forms have been collected and analyzed, profiles of student writers and conferencing activities (writing problems most frequently cited by students, tutors' perceptions of most common problems, most frequent types of writing done by students, strategies most often employed by tutors) can be drawn up.

Validity, Reliability, and Normative Data: The developers of WCTRF do not report any data of this type.

Ordering Information: ED 236 631

Related Documents:

Luban, Nina, Ann Matsuhashi, and Tom Reigstad. "One-to-One to Write: Establishing an Individual-Conference Writing Place." *English Journal* 67 (November 1978): 30–35. (EJ 190 310)

Reigstad, Tom. "The Writing Place at Buffalo: The First Years." *ADE Bulletin* (September 1980): 31–35. (EJ 234 080)

Reigstad, Tom, Michael Williamson, and Ann Matsuhashi. "The Buffalo City Schools' 'Writing Place' Project: Evaluating the Effectiveness of a Supplemental Tutorial Writing Program." *English Record* (Spring 1980): 2–9.

Category: Writing

Title: Questionnaire for Identifying Writer's Block (QIWB)

Author: Mike Rose

Age Range: Senior high and postsecondary-adult

Description of Instrument:

Purpose: To identify students with writer's block.

Date of Construction: 1981

Physical Description: QIWB is an attitude questionnaire of twenty-four items separated into five subscales: "Blocking" (Items 7, 9, 12, 16, 17, 22, 24), "Lateness" (4, 14), "Premature Editing" (3, 8, 18), "Strategies for Complexity" (5, 11, 15, 19, 23), and "Attitudes" (1, 2, 6, 10, 13, 20, 21). If a teacher or a researcher wishes simply to identify blockers, he or she can administer only the items within the behavioral subscales "Blocking" and "Lateness." If further cognitive diagnosis is desired, then the items in the cognitive/behavioral and cognitive/attitudinal subscales can also be administered.

Validity, Reliability, and Normative Data: An assumption of QIWB is that writer's block is not solely an emotional response to fear of evaluation or self-revelation but instead reflects, at least partly, cognitive difficulties, such as a lack of strategies for composing complex discourse or a tendency to edit inappropriately early in the composing process.

Reliability coefficients for the five subscales ranged from .72 to .87 with a median coefficient of .84.

Since the study's conceptualization of writer's block posited that "Lateness," "Premature Editing," "Strategies for Complexity," and "Attitudes" share a positive relation to "Blocking," but measure different aspects and manifestations of it, a pattern of moderate correlations should have emerged among subscales. This pattern resulted: correlations ranged from .37 to .59 with a median correlation of 40.5.

Given the study's conceptualization of writer's block, some percentage of response variance on the "Blocking" subscale (the primary behavioral indicator of writer's block) should be predicted from each of the other subscales. A regression analysis demonstrated that the subscales "Lateness," "Premature Editing," "Strategies for Complexity," and "Attitudes" predicted 52 percent of the variance in "Blocking" response.

One test of validity was conducted by comparing the questionnaire responses of an (admittedly small) subsample of ten chosen for the study's stimulated recall investigation with that subsample's subsequent comments and behaviors during stimulated recall interviews (see Rose 1981). These students' comments and behaviors almost uniformly supported earlier questionnaire responses. In the handful of instances where comments and behaviors seemed to contradict earlier responses, further investigation removed or explained the contradiction.

Ordering Information: ED 236 652

Related Documents:

Rose, Mike. "Rigid Rules, Inflexible Plans, and the Stifling of Language: A Cognitive Analysis of Writer's Block." *College Composition and Communication* 31 (1980): 389–401. (EJ 240 357)

———. "The Cognitive Dimension of Writer's Block: An Examination of University Students." Ph.D. diss., University of California, Los Angeles, 1981.

———. *Writer's Block: The Cognitive Dimension.* Carbondale, Ill.: Southern Illinois University Press, 1984.

Category: Writing

Title: Analyzing the Structure of Children's Compositions

Author: Abraham Stahl

Age Range: Grades 2 through 8

Description of Instrument:

Purpose: To study the development of structure in children's compositions—the differences between children of different age, sex, and socioeconomic status.

Date of Construction: 1968

Physical Description: This instrument attends to structural units larger than sentences—to the overall organization of compositions. It was designed for the analysis of factual-descriptive compositions. (Subjects responded to this topic: "Write a description of your home, its rooms and their contents in such a way that someone who has not visited it can form an idea of how it looks.") The instrument includes nine categories, each containing five types of structure. The types are marked in descending numerical order from 5 to 1. Type number 5 is considered the best from the point of view of structure, 4 is less good, and so forth.

A. Indicated Order
   5. Implicit and Consistent
   4. Explicit and Consistent
   3. Absence of Indicated Order
   2. Implicit and Inconsistent
   1. Explicit and Inconsistent

B. Principle of Selection
   5. General: Concerned with Essentials
   4. General: Concerned with Essentials but also Includes Details
   3. Inconsistent
   2. Omission of Essentials
   1. Absence of any Conscious Choice

C. Methods of Arrangement
   5. Comprehensive
   4. Surveying
   3. Associative
   2. Egocentric
   1. Enumerating

D. Syntax
   5. Connection Between Sentences
   4. "Complex" Sentences
   3. "Simple" Sentences
   2. Incomplete Sentences
   1. Words or Phrases

- E. Balance
    - 5. Well-Balanced
    - 4. Prominence Given to an Important Part of the Subject
    - 3. Prominence Given to Matter of Secondary Importance
    - 2. Lack of Balance
    - 1. Digression from the Subject
- F. Organization
    - 5. Advanced Planning
    - 4. Structural Correction
    - 3. Absence of Signs of Planning or Structural Correction
    - 2. Faulty Organization
    - 1. Complete Obscurity
- G. Connectives
    - 5. Sophisticated Use
    - 4. Normal and Correct Use
    - 3. Poor Use
    - 2. Faulty Use
    - 1. Complete Absence of Connectives
- H. Opening
    - 5. Introductory Opening
    - 4. "Surveying" Opening
    - 3. "Inclusive" Opening
    - 2. Absence of Opening, but Begins with an Important Matter
    - 1. Absence of Opening; Begins with an Insignificant Detail
- I. Conclusion
    - 5. Formal Termination
    - 4. Announcement of Ending
    - 3. End of Description
    - 2. Subjective Ending
    - 1. Lack of Conclusion

Each category and type of structure is defined, and scoring guidelines and examples are provided for all forty-five possible structural ratings. For example, Category A, Structure Type 5 follows:

> The category of *indicated order* deals with the degree of consistency to which the composition is arranged in relation to its opening passage.
>
> 5. *Implicit and Consistent.* The composition opens with a list of rooms, then describes their contents in precisely the same, or in precisely the reverse, order.
>
> *Notes:* A. A general opening statement: "Our home has 4 rooms, a kitchen and a bathroom" without designating the rooms, relegates the composition to type 3 of this category.
>
> B. "Precisely the reverse order"—applies only where there are more than 2 rooms.
>
> *Example:* ". . . there are three rooms: one room for the children, one room for guests and one room for the

> parents. In the children's room . . . In the guest room . . . In the parents' room . . ."
>
> It should be noted that the list does not always come at the very beginning.

Raters are encouraged to proceed in the order in which the instrument is arranged and not to consider related categories simultaneously so they may arrive at an independent judgment of each category.

Validity, Reliability, and Normative Data: The instrument was used with a sample of 400 Israeli children in grades 2, 5, 8, and 11. (Grade 11 students found the topic "childish.") Five hundred twenty-two compositions were analyzed twice by the author. In 841 of 4,698 separate judgments (522 × 9 categories), categorizations were identical. In case of differences, compositions were read a third time. Because the author could find no related instrument in the literature, he was unable to put the compositions to another test independent of his.

Ordering Information: See related documents.

Related Documents:

Stahl, Abraham. "Structural Analysis of Children's Compositions." *Research in the Teaching of English* 8 (1974): 184–205. (EJ 169 457)

———. "The Structure of Children's Compositions: Developmental and Ethnic Differences." *Research in the Teaching of English* 11 (1977): 156–63. (EJ 105 716)

Category: Writing

Title: Measures of a Writer's Choice (MWC)

Authors: Sauli Takala
Alan C. Purves
Annette Buckmaster

Age Range: Junior high through postsecondary-adult

Description of Instrument:

Purpose: To indicate the probability that a writer would make particular decisions about audience, topic, structure, and style.

Date of Construction: 1982

Physical Description: MWC presents a series of statements about writers' decisions. The statements are grouped under presumed major categories of decisions writers must make as they compose: audience, topic, tone, development, structure, and style. The student indicates how likely he or she would be to make each decision just the way it is stated. MWC is reproduced on the following page.

The authors of MWC suggest that "Alternate phrasings to this kind of questionnaire might ask the same questions in terms of 'To get a high mark an essay should . . .'; 'I would like to be able to write an essay that would . . .'; 'My teachers would like an essay that. . . .'"

Validity, Reliability, and Normative Data: MWC at present is only a proposal from an international group studying national differences in writing style (Purves and Takala 1982). It is proposed only as a "surrogate" for actual writing, an indirect measure that might still yield information of use to evaluators. Even in its present limited form, MWC provides a model that evaluators could adapt in many ways for their own purposes. In this format, statements could be written that would reflect specific objectives of a writing program, especially those objectives concerned with the writing process—with planning, goal-setting, drafting, and revising.

Ordering Information: Full instrument reproduced above.

Related Document:

Purves, A. C., and S. Takala, eds. "An International Perspective on the Evaluation of Written Composition." *Evaluation in Education* 5 (1982).

Measure of a Writer's Choice

As you know, when a person begins to write an essay or a paper on a given topic, that person has to make some decisions. Different writers make different decisions. We would like to find out which decisions you think are important and what decisions you would make.

You have been given an assignment to write a short essay of two to three pages on the topic "What Is a Friend?" This topic is part of an examination. Below is a number of decisions that students have made. Please rate each decision by putting an *X* in the space by the decision you would probably make.

	I would do this	*I would probably do this*	*The choice is not important*	*I would probably not do this*	*I would certainly not do this*
[Audience]					
1. I would choose a particular person, a friend or a teacher, and pretend that I was writing to that person.	______	______	______	______	______
2. I would write a first draft and then recopy it to make sure the examiner would find no mistakes.	______	______	______	______	______
[Topic Selection]					
3. I would think about the topic and select one or at most two points and concentrate on those.	______	______	______	______	______
4. I would make a list of as many different kinds of friendship and types of friends and try to include them all in my essay.	______	______	______	______	______
[Personal-Impersonal]					
5. I would begin by writing about my friends and use my feelings as an example.	______	______	______	______	______

	I would do this	*I would probably do this*	*The choice is not important*	*I would probably not do this*	*I would certainly not do this*
6. I would try to show the reader what I really think about friends.	______	______	______	______	______
[Abstract-Concrete]					
7. I would try to make the essay show the reader what a good general definition would be.	______	______	______	______	______
8. I would use many different examples of real people or people in stories to illustrate my point.	______	______	______	______	______
[Structural]					
9. I would arrange my essay to show the examiner that I can think logically.	______	______	______	______	______
10. I would try to have the essay show how my mind actually works with thoughts, feelings, images, and examples as they come up.	______	______	______	______	______
[Stylistic]					
11. I would try to keep the words as simple as possible and say things without any metaphors, similes, or other tricks that writers use.	______	______	______	______	______
12. I would use words and images to try to make the reader see friends or feel about friends as I do.	______	______	______	______	______

Category: Writing

Title: Unipolar Scale for Evaluating Writing (USEW)

Authors: Sauli Takala
Alan C. Purves
Annette Buckmaster

Age Range: Intermediate through postsecondary–adult

Description of Instrument:

Purpose: To contrast the standards for evaluating writing of different groups of raters.

Date of Construction: 1982

Physical Description: USEW is a set of unipolar semantic differential scales for evaluating essays. The rater considers whether the essay is low or high in thirty different qualities of writing. For evaluation projects or comparison group research projects, the mean score could be reported for each scale. For studies of style, the results of USEW could be factor-analyzed to identify characteristic features of groups of essays or evaluation standards of groups of raters. USEW is reproduced below.

Unipolar Scales for Evaluating Writing

Please read the attached essay and then give your opinion of the essay by placing an X in the space after each adjective that is closest to your judgment of the essay.

____________ Essay Number

____________ Rater Number

	Low in this Quality						High in this Quality
Detached	____ 1	____ 2	____ 3	____ 4	____ 5	____ 6	____ 7
Good	____ 1	____ 2	____ 3	____ 4	____ 5	____ 6	____ 7
Original	____ 1	____ 2	____ 3	____ 4	____ 5	____ 6	____ 7
Detailed	____ 1	____ 2	____ 3	____ 4	____ 5	____ 6	____ 7
Concrete	____ 1	____ 2	____ 3	____ 4	____ 5	____ 6	____ 7
Interesting	____ 1	____ 2	____ 3	____ 4	____ 5	____ 6	____ 7
Metaphoric	____ 1	____ 2	____ 3	____ 4	____ 5	____ 6	____ 7

Clear	1	2	3	4	5	6	7
Selective	1	2	3	4	5	6	7
Plain	1	2	3	4	5	6	7
Vivid	1	2	3	4	5	6	7
Concise	1	2	3	4	5	6	7
Coherent	1	2	3	4	5	6	7
Direct	1	2	3	4	5	6	7
Personal	1	2	3	4	5	6	7
Particular	1	2	3	4	5	6	7
Abstract	1	2	3	4	5	6	7
Elaborate	1	2	3	4	5	6	7
Unified	1	2	3	4	5	6	7
Literal	1	2	3	4	5	6	7
Logical	1	2	3	4	5	6	7
General	1	2	3	4	5	6	7
Critical	1	2	3	4	5	6	7
Flowing	1	2	3	4	5	6	7
Sensuous	1	2	3	4	5	6	7
Artificial	1	2	3	4	5	6	7
Smooth	1	2	3	4	5	6	7
Individual	1	2	3	4	5	6	7
Vague	1	2	3	4	5	6	7
Associative	1	2	3	4	5	6	7

Validity, Reliability, and Normative Data: In the version reproduced above, USEW is only a proposal from an international group studying national differences in writing style (Purves and Takala 1982). It is based, however, on an established tradition of empirical research on judgments of complex entities like literary works or students' essays.

Ordering Information: Full instrument reproduced above.

Related Document:

Purves, A. C., and S. Takala, eds. "An International Perspective on the Evaluation of Written Composition." *Evaluation in Education* 5 (1982).

Category: Writing

Title: Models for Analysis of Writing (MAW)

Authors: Andrew Wilkinson
Gillian Barnsley
Peter Hanna
Margaret Swan

Age Range: Primary through senior high

Description of Instrument:

Purpose: To describe and evaluate the characteristics of student writing.

Date of Construction: 1980

Physical Description: MAW is actually four different systems of analysis, each based on a "model" or hypothesis about writing development between ages seven and thirteen. The authors describe the four models this way:

> *Cognitive.* The basis of this model is a movement from an undifferentiated world to a world organized by the mind, from a world of instances to a world related by generalities and abstractions.
>
> *Affective.* Development is seen as being in three movements—one towards a greater awareness of self, a second towards a greater awareness of neighbour as self, a third towards an inter-engagement of reality and imagination.
>
> *Moral.* 'Anomy' or lawlessness gives way to 'heteronomy' or rule by fear of punishment, which in turn gives way to 'socionomy' or rule by a sense of reciprocity with others which finally leads to the emergence of 'autonomy' or self-rule.
>
> *Stylistic.* Development is seen as choices in relation to a norm of the simple, literal, affirmative sentence which characterizes children's early writing. Features, such as structure, cohesion, verbal competence, syntax, reader awareness, sense of appropriateness, undergo modification. (3, 4)

The Cognitive Model offers criteria for identifying and evaluating description, interpretation, generalization, and speculation. The Affective Model includes criteria for awareness of self, other people, one's readers, environment, and reality. The Moral Model provides criteria for evaluating whether in their writing students judge themselves and others by physical characteristics and consequences, in terms of punishments or rewards, according to the status quo, or in terms of conventional norms and rules, intentions or motives, abstract concepts, or a personally developed value system. The Stylistic Model includes criteria for syntax, vocabulary, organization, cohesion, awareness of readers, appropriateness, and effectiveness. Under this last aspect of style—effectiveness—are separate sets of criteria for evaluating autobiography, narrative, explanation, and argument.

The writers offer this clarification about MAW: "These models are not intended to be used as day-to-day marking schemes; but to heighten levels of awareness. Their detail enables them to pay due regard to the varieties of activity going on in the process of writing. In one sense they are assessment instruments, but only in the sense that assessment is an essential part of education. . . ."

Here are brief sections from each of the four models in MAW:

> ### Cognitive Model
>
> *Generalizing*
>
> Abstracting—using abstract terms as well as concrete ones e.g., "People say children should go to school," "The *players* move alternately, white beginning . . ."
>
> Summarizing—e.g., "So you see Topcat won," "The object of owning property is to collect rents from opponents stopping there," "The first person to do that is the winner."
>
> Overall Evaluation—e.g., "So Topcat won by being more clever," "The main object of the game is to meld seven cards of a kind."
>
> Concluding—e.g., "So he decided never to enter the race again," "These seven points show just how ludicrous that suggestion really was."
>
> Reflecting—generalizing with reference to external rules or principles, e.g., "This phase would generally have lasted several years."
>
> Classifying—links between generalizations sustained in a classificatory system. (72)
>
> ### Affective Model
>
> *Other People*
>
> The writer shows an awareness of others both in relation to himself or herself and as distinct identities.
>
> —records the mere existence of other people as having been present. This is the single dimension: others are present—acting, speaking—but no emotion is apparent by inference, e.g., "The two boys went for a walk with their mother and they got lost and they came to a fence and that fence was electric and they was not lost. . . ."
>
> —begins to indicate the separateness of others by, e.g., giving their actual words or significant actions. "I woke up, had my breakfast" is probably not significant; "the old man smiled" may well be.
>
> —the thoughts and feelings of others by quotation of actual words, perhaps as a dialogue, or by description of them, or

actions indicating them. More perception called for than in the previous category though it might be fairly conventional.

Analytical, interpretative comments on aspects of character and behaviour; or insightful quotation or dialogue.

Consistently realized presentation of another person by a variety of means, perhaps by assuming persona.

Ability to see a person and his or her interactions in extended context (e.g., a character in a novel). (73, 74)

Moral Model

Attitudes/judgements about self/others and events.

Judging self/others by physical characteristics or consequences, e.g., "She was ugly, so she was bad." "He broke fifteen cups—naughty." Judging events by pain-pleasure to the self, e.g., "It was a bad day. I hurt my hand." "It was a good birthday. I got lots of presents." "A bad accident—the fence was smashed up." Principle of self-gratification—"anomy."

Judging self/others and events in terms of punishments/rewards. "I won't do that, Mummy will hit me." "I'll tell Daddy on you and he will beat you up." "If I do the dishes, Mummy will give me a new bat." Events judged as rewards/punishments, e.g., "I must have been naughty last night, the fridge hit me." Heteronomy.

Judging self/others according to the status quo. Mother, father, teacher, policeman good by right of status; the wicked witch, the evil step-father bad by right of convention, e.g., "I hated the Jerries, I used to call them stupid idiots." Reciprocity restricted to the child's immediate circle, e.g., "I won't do that—it will upset Mummy." Social approval/disapproval internalized in terms of whether behaviour upsets others or not. Stereotypic thinking. Events judged in terms of effects on other *people*. "It was a bad accident. All the passengers were badly hurt." Socionomy (internal).

Judging self/others in terms of conventional norms/rules, e.g., "It's wrong to steal. It is against the law." Conformist orientation. Rules are applied literally on the principle of equity or fairness. "It's not fair. We all did it, so John should be punished the same as us. We all broke the rule." Socionomy (external).

Judging self/others in terms of intention or motive, regardless of status or power, e.g., "She didn't mean to drop those plates, so she shouldn't be punished." "Teacher was wrong, because she punished all of us instead of finding out who did it."

Judging self/others in terms of abstract concepts such as a universal respect for the individual rather than in terms of conventional norms of right/wrong conduct. The morality of individual conscience. Rules seen as arbitrary and changeable. Autonomy.

Judgement of self/others in terms of a personally developed value system. (75, 76)

Stylistic Model

Cohesion
Cohesive devices are employed to maintain continuity between one part of the text and another. Just as grammar establishes the structural relationship within clause or sentence, so cohesion establishes the semantic relationship within the text. There is development from the relatively unrelated to the fully related parts in a text.

Few cohesive devices employed effectively. Pronouns, where used, sometimes have no specific referent or are used imprecisely. Ellipsis, when employed, often shows no clear understanding of the referent, e.g., "If they miss [the goal?] the other player has his or her [turn?] !" Little lexical cohesion. Most common conjunctions: "and," "so," "then."

Marked increase in cohesive devices. Sequential and concluding conjunctions, e.g., "afterwards," "finally," "eventually."

Use of temporal conjunctions, e.g., "when," "first," "first of all."

Use of causal conjunctions, e.g., "so," "because."

Use of "but" in an adversative/contrastive way.

Some use of demonstratives as adverbs of place, e.g., "here," "there."

Some substitution, e.g., "one," "other," "some."

Nominal substitution, e.g., "one," "the same" and verbal substitution, e.g., "do so," "be so."

Appearance of low-level general terms, e.g., "people," "things."

Greater awareness of textual coherence to clarify and define meaning. Emphatic cohesive conjunctions, e.g., "too," "even," "also." Use of comparatives, e.g., "identical," "similar," "more," "less," and superlatives, e.g., "the wealthiest."

Development of logical coherence. Use of superordinates. A wider range of adversatives employed, e.g., "however," "on the other hand," "though."

A wide range of cohesive devices employed. For example, reiteration, synonyms, antonyms, parallelism, contrast, assonance, alliteration, echoic words, etc. (78, 79)

Validity, Reliability, and Normative Data: Since MAW was used only to guide a study of the writing development of children between the ages of seven and thirteen, not to score writing samples, it produced no reliability or normative data. The study reporting the use of MAW does provide a strong characterization of the writing of British children ages seven, ten, and thirteen. The validity of MAW as a guide to describing writing comes from its grounding in current theories of cognitive and moral development and in research on language development. Though it provides no scoring instruments in its present form, various parts of MAW could be adapted for different evaluation purposes.

Ordering Information: See related document.

Related Document:

Wilkinson, A., G. Barnsley, P. Hanna, and M. Swan. *Assessing Language Development*. New York: Oxford University Press, 1980.

Category: Writing

Title: Evaluating Course and Teacher Effectiveness (ECTE)

Authors: Stephen P. Witte
John A. Daly
Lester Faigley
William R. Koch

Age Range: Senior high and postsecondary-adult

Description of Instrument:

Purpose: To evaluate the effectiveness of writing courses through student perceptions of the instructor and the course.

Date of Construction: 1981

Physical Description: ECTE is a Likert-type instrument in two forms: a short form of twenty items and a long form of ninety-nine items. In the long form, Part I assesses students' perceptions of the teacher, teaching methods, and value of the course. Part II assesses nine different factors of writing instruction. The short form focuses on teacher effectiveness and course content.

Students respond to each item on a 5-point Likert scale: strongly agree, agree, undecided, disagree, strongly disagree, with a numerical value range of 1–5.

The short form of ECTE is reproduced below without the Likert scale for each item.

Short Form of ECTE

1. My teacher in this course is a very good one.
2. My teacher in this course is one of the most helpful teachers I have ever had.
3. The instructor is intellectually stimulating.
4. My work as a student in this course is evaluated fairly.
5. The instructor does a good job of teaching students how to write different types of papers.
6. The instructor is good about teaching students how to argue a thesis or position.
7. The instructor puts too much emphasis on "correct" grammar in student writing.
8. The instructor clearly connects what he or she teaches in class with the writing assignments he or she makes.
9. The instructor is good about teaching me to evaluate my own papers so that I can better revise them.

10. The instructor is good about teaching me to consider the needs of my audience when I am thinking about a writing topic.
11. The instructor is good about using class time to help me as I am writing my papers.
12. The instructor spends too much time on some things and not enough time on others.
13. In evaluating my writing, the instructor uses standards that are too high.
14. The instructor's comments on my papers are easy to understand.
15. The instructor is good about teaching me how to support main ideas in my papers through examples and specific details.
16. The instructor does a good job of using examples of writing in teaching students how to write.
17. The instructor is good about encouraging me to join in class discussions.
18. The instructor is good about trying to increase my confidence about writing.
19. What this course teaches is very useful to the student right now.
20. What I am learning from this course is valuable.
21. What this course teaches will be useful to its students in the future.

Validity, Reliability, and Normative Data: ECTE was developed in four stages: (1) its developers collected a large pool of statements about writing teachers and courses; (2) they administered pilot questionnaires and generated an inter-item correlation matrix in order to reduce the number of items and revise the wording of items retained; (3) they administered the revised instrument to a large number of students in different colleges and factor-analyzed the inter-item correlation matrix generated from student responses; and (4) they modified the instrument still further, readministered it, factor-analyzed again, and then selected items for the final instrument. In the fourth stage, 1,552 students from nine colleges responded to the instrument.

The coefficient alpha reliability of Part I of the long form is .93. For the nine factors in Part II reliability ranges from .70 to .93. The reliability of the short form is .86.

Ordering Information: ED 211 981

Related Document:

Witte, S. P., J. A. Daly, L. Faigley, and W. R. Koch. *The Empirical Development of an Instrument for Reporting Course and Teacher Effectiveness in College Writing Classes.* Austin: University of Texas Writing Program Assessment Project, 1981. (ED 211 981)

Miscellaneous

Category: Miscellaneous

Title: Film Preference Instrument

Author: Carole A. Cox

Age Range: Grades 4 and 5

Description of Instrument:

Purpose: To measure the interest of fourth- and fifth-grade children in certain films in order to determine whether or not an underlying pattern of children's film preferences exists.

Date of Construction: 1974

Physical Description: The directions for administration and scoring, shown below, provide a description of this instrument:

> *Directions:* (read aloud to children)
>
> Today you are going to see some (more) films. I would like to know how much boys and girls like or dislike certain films.
>
> This Is Not a Test. So please do not say you like a film because you think you should, or adults think you should. I want to know how you honestly feel about each film because I don't really know and I think it is important that we know what boys and girls really like or don't like in films.
>
> Please Look at Your Questionnaire. Fill in your name and check boy or girl. You can show how you feel about a film by putting a circle around one of the five sentences.
>
> Look at the top sentence. Who can read it for me? That's right. "I didn't like it at all." You would circle that sentence if you didn't like a film at all. Look at the sentence next to it. Who can read it for me? That's right. That means if you circle "I didn't like it at all" you would rather have been doing something else than watching the film. It was boring, you were tired of it, or thinking of something else.
>
> Now, can anyone give me an example of a film you have seen in school, on TV, or in a movie theatre, or a TV show that you would say, "I didn't like it at all." You would have left or done something else if you could have.
>
> Look at the next sentence. Who can read it for me? (etc., for all the sentences.)
>
> After each film, you will circle one of the sentences.
>
> Remember, Please Show How You Honestly Feel about Each Film. Do not consider what I might think, or your teacher, or your parents, or your friends. There are not right or wrong answers. You may feel differently about each film. You will like some films and not like other films. You may circle a different sentence for

each film. That is fine. Everyone has likes and dislikes and everyone likes some films less or more than others.

Remember: I want to know how much each one of you really likes or dislikes each film. Questions?

Name	Boy	Girl	Grade	School
________	_____	_____	________	________

How much did you like this film?

Film ______(title)______

I didn't like it at all.	(I would rather have done something else.)
I didn't like it very much.	(I wouldn't want to see it again.)
It was o.k.	(I wouldn't mind seeing it again.)
It was good.	(I would like to see it again.)
It was great!	(I could see it many times without getting tired of it.)

The ratings for each film were weighted as follows:

1 = I didn't like it at all.
2 = I didn't like it very much.
3 = It was o.k.
4 = It was good.
5 = It was great!

These scores were used to compute mean ratings later used for factor analysis.

Validity, Reliability, and Normative Data: The most important consideration in the creation of the questionnaire was the need to produce an instrument that would measure a range of responses to twenty-four films. The goal was to obtain data upon which factor analytic procedures could be used to determine the film preference patterns of fourth- and fifth-grade children as they pertain to the content and form of the short film. Results of a first pilot study underscored the importance of this need. Children rated all films nearly the same even though the films varied greatly in content and form. A new questionnaire was tested in a second pilot study with fifty-five fourth- and fifth-grade children. Since it showed a range of mean film ratings from 1.89 to 3.88 on a scale of 1 to 5, it was used in the major investigation. Further evidence of the usefulness of this instrument was provided when 344 fourth- and fifth-grade children in two elementary schools in Baton Rouge, Louisiana, viewed and rated twenty-four films over a six-week period. Mean film ratings ranged from 1.75 to 4.67 on a scale of 1 to 5. In order to transform the set of twenty-four variables, or films in this study, into a new set of composite variables, the film ratings of the 344 children were subjected to principal components factor analysis with varimax rotations. The eight factors which emerged are shown on the following page with mean scores for each factor.

Rank	Factor	Average of Film Means
1	Real Children/Work and Play	4.57
2	Children and People/Suspense	4.39
3	Fantasy/Excitement	4.01
4	Action/Sport	3.90
5	Fantasy/Humor	3.49
6	Animals/Humor	3.04
7	Nature	2.58
8	Abstract Visual	2.09

Film rating scale: 1—I didn't like it at all; 2—I didn't like it very much; 3--It was o.k.; 4—It was good; 5—It was great!

Ordering Information: See related document.

Related Document:

Cox, Carole A. "Film Preference Patterns of Fourth and Fifth Grade Children." Ph.D. diss., University of Minnesota, 1975.

Appendix

Category Index to *Measures for Research and Evaluation in the English Language Arts, Volume 1*

In this index the measures are listed by category or type of instrument, arranged alphabetically by category heading: Language Development, Listening, Literature, Reading, Standard English as a Second Language or Dialect, Teacher Competency, and Writing. The final category is Miscellaneous. Within each category measures are listed alphabetically by the first author's name.

Cross indexing is also provided for the age range appropriateness of each instrument. Following is a key to age range headings:

PS	pre-school (birth to K)
P	primary (Grades 1-3)
I	intermediate (Grades 4-6)
JH	junior high (Grades 7-9)
SH	senior high
PS-A	postsecondary-adult

Language Development

Author	*Title*	*Age Range*	*Page*
Bentler, P.M. Lavoie, A.L.	A Nonverbal Semantic Differential	Unspecified	3
Berko, J.	Berko's Test of Morphology	PS,P	5
Calvert, K.	K-Ratio Index	I	9
Chomsky, C.	Acquisition of Syntax Experiments	PS,P,I	11
Fisher, C.J.	Linguistic Structures Repetition Test	P	13
Fraser, C. Bellugi, U. Brown, R.	The Imitation-Comprehension-Production Test	PS	15
Fryburg, E.L.	The Test of Cognition	P	18
Golub, L.S. Fredrick, W.C.	Linguistic Ability Measurement Program	I,JH	23
Golub, L.S. Fredrick, W.C. Nelson, N.J. Frayer, D.A. Harris, M.L.	Schema for Testing Language Arts Concept Attainment	I	27

Koziol, Jr., S.M.	Noun Plural Development Test	PS,P	31
McNamara, T.C. Ayrer, J.E. Farber, I.J.	Semantic Differential Scales for Use with Inner-City Pupils	I,JH	34
Marcus, A.D.	A Test of Sentence Meaning	I,JH	36
O'Donnell, R.C.	Perception of Alternate Structures Test	JH,SH	39
Slobin, D.I. Welsh, C.A.	Elicited Imitation	PS	41

Listening

Author	*Title*	*Age Range*	*Page*
Bowdidge, J.S.	Cloze Listening Test	JH,SH	47
Sigelman, C.K.	Giving and Taking Directions	JH,SH	49

Literature

Author	*Title*	*Age Range*	*Page*
Andresen, O.	Literary Profundity Test	SH,PS-A	55
Beaven, M.H.	Questionnaire: Responses to Feminine Characters in Literature	SH,PS-A	56
Benson, L.M.	Describing and Evaluating Classroom Discussions of Poems	JH,SH,PS-A	58
Carnegie Curriculum Study Center in English Staff	Literary Discernment Test	SH,PS-A	59
	Literary Preference Questionnaire	SH,PS-A	61
Foreman, E.	Literary Appreciation of Adolescents	JH,SH,PS-A	63
Harpin, W.S.	Literary Discrimination Test	SH,PS-A	65
Purves, A. International Association for the Evaluation of Educational Achievement, Committee for Literature	Achievement Measures (Understanding and Interpreting Short Stories)	JH,SH,PS-A	66
	Literary Transfer and Interest in Reading Literature	JH,SH,PS-A	68
	Response Preference Measure	JH,SH,PS-A	73

Rees, R.D. Pedersen, D.M.	Poetic Evaluation Rating Scale	JH,SH,PS–A	75
Reich, C.	Novel Reading Maturity Scale	JH,SH,PS–A	77
Sanders, P.L.	Criteria for the Evaluation of Free-Response Interpretive Essays	JH,SH,PS–A	80
Wall, E.R.	Topical Analysis of the Content of Literature Discussions	JH,SH,PS–A	180
Zais, R.	Sophistication of Reading Interests Scale	JH,SH	81

Reading

Author	*Title*	*Age Range*	*Page*
Bickley, A.C. Bickley, R.T. Cowart, H.	Oral Paradigmatic/Syntagmatic Language Inventory	I	85
Botel, M. Granowsky, A.	Syntactic Complexity Formula	P,I	87
Burt, H.R.	Similes Test	I	90
Cosens, G.V.	Cloze Test for Deletion Produced Structures	P	92
DeLancey, R.W.	Recognition of Linguistic Structures	I,JH	95
Gavin, Sister Mary Therese	S–N Auditory Discrimination Test	PS,P	97
Golub, L.S.	Syntactic Density Score	P,I,JH,SH,PS–A	99
Jewell, C.B.	The Jewell High School Reading Comprehension Test	SH	101
Kretschmer, J.C.	Experimental Test of Piagetian Concrete Operations in a Reading Format	P,I	103
Little, P.S.	Sentence Interpretation Test	I	106
Miller, L.A.	Tests of Anaphoric Reference-Multiple Choice Format; Tests of Anaphoric Reference-Cloze Format	P	110
Montague, M.	The Contextual Ambiguity Test	I,JH	112
Peltz, F.K.	Linguistic Analysis Worksheet	SH	115
Powell, A.	Primary Reading Attitude Index	P	118
	Intermediate Reading Attitude Index	I	119

Robertson, J.E.	Connective Reading Test; Written Connectives Test	I	121
Rystrom, R.	Rystrom Reading Comprehension Test	I	123
Shepherd, J.F.	Morpheme Knowledge Test	SH,PS–A	125
Simons, H.D.	Deep Structure Recovery Test	I	127
Stennett, R.G. Smythe, P.C. Hardy, M. Pinkney, J. Fairbairn, A.	Developmental Patterns in Elemental Reading Skills	PS,P	128
Taylor, W.L.	Cloze Procedure	P,I,JH,SH,PS–A	139
Twohig, B.T.	Letter Directionality Test; Word Directionality Test; Sentence Directionality Test	PS,P	141
Wright, O.T.	Identification of Simple and Compound Vowels by First Graders	P	143

Standard English as a Second Language or Dialect

Author	*Title*	*Age Range*	*Page*
Beardsmore, H.B. Renkin, A.	A Test of Spoken English	PS–A	147
Brengelman, F.H. Manning, J.C.	Linguistic Capacity Index	P	150
Jersey City State College; Vineland School District; National Consortia for Bilingual Education	Parent Questionnaire on Bilingual Education	PS–A	153
Loban, W.	Categories for Tallying Problems in Oral Language	PS,P,I,JH	155
Mantell, A.	Test of Language Judgment	I	157
McNamara, T.C. Ayrer, J.E. Farber, I.J.	Semantic Differential Scales for Use with Inner-City Pupils	I,JH	34
Natalicio, D.S. Williams, F.	Questionnaire for Evaluators of Black Language Samples; Questionnaire for Evaluators of Mexican-American Language Samples	PS,P	159

Politzer, R.L. Hoover, M.R.	Standard Discrimination Test	P,I	161
Rystrom, R.C.	Rystrom Dialect Test	PS,P	165

Teacher Competency

Author	*Title*	*Age Range*	*Page*
Benson, L.M.	Describing and Evaluating Classroom Discussions of Poems	JH,SH,PS–A	58
Gallo, D.R.	Poetry Methods Rating Scale	Unspecified	171
Hipple, T.W. Giblin, T.R.	Professional Reading of Teachers Questionnaire	Unspecified	173
Hoetker, J. Robb, R.	The "Place of Drama" Questionnaire	Unspecified	174
Madsen, A.L.	Instrument to Survey Knowledge of Literary Criticism	Unspecified	175
National Council of Teachers of English, Commission on Composition	Composition Opinionnaire: The Student's Right to Write	Unspecified	177
Smith, V.H.	Composition Rating Scale	Unspecified	178
Wall, E.R.	Topical Analysis of the Content of Literature Discussions	JH,SH,PS–A	180

Writing

Author	*Title*	*Age Range*	*Page*
Ashida, M.E.	A Standard Corpus of Contemporary American Expository Essays	SH,PS–A	185
California Association of Teachers of English, Joint Sub-Committee on Composition	A Scale for Evaluation of High School Student Essays	JH,SH	187
Dauterman, F.	Syntactic Maturity Test for Narrative Writing	I,JH,SH,PS–A	188
Diederich, P. French, J. Carlton, S.	E.T.S. Composition Evaluation Scales	I,JH,SH,PS–A	190
Dixon, E.	Indexes of Syntactic Maturity (Dixon-Hunt-Christensen)	P,I,JH,SH,PS–A	191

Dupuis, M.M.	Transformational Analysis of Compositions	I,JH,SH,PS-A	193
Glazer, J.	Glazer Narrative Composition Scale	P,I,JH	195
Golub, L.S.	Syntactic Density Score	P,I,JH,SH,PS-A	99
London Association for the Teaching of English, Sub-Committee on Assessing Composition	Evaluation Scale for Personal Writing	SH	197
Mellon, J.	Factors of Syntactic Fluency (Hunt-Mellon)	P,I,JH,SH,PS-A	199
O'Donnell, R.C. Hunt, K.W.	Syntactic Maturity Test	P,I,JH,SH,PS-A	201
Sager, C.	Sager Writing Scale	I,JH	203
Schroeder, T.S.	Schroeder Composition Scale	P,I,JH	204
Tway, E.	Literary Rating Scale	P,I,JH	205

Miscellaneous

Author	*Title*	*Age Range*	*Page*
Evanechko, P.O.	Semantic Features Test	I,JH	209
Moore, B.J.	The Heinsen-Moore Test of Visual Processing Skills	P	212

Author Index

In this index the measures are listed alphabetically by the author's last name.

Author	*Title*	*Page*
Abbott, Robert T.	jt. auth., see Schuessler, Brian F.	
Alparaque, Idrenne	Assessment of Instructional Terms	67
Anderson, Linda	jt. auth., see Duffy, Gerald	
Askov, Eunice N.	jt. auth., see Dupuis, Mary M.	
Baker, Christopher J.	jt. auth., see Hoffman, James V.	
Barnsley, Gillian	jt. auth., see Wilkinson, Andrew	
Benterud, Julianna G.	Categories for Observing Language Arts Instruction	135
Black, Janet K.	Interactional Competency Checklist	7
Blake, Howard E.	Diagnosis of Language Competency Inventory	10
Brause, Rita S.	Comprehension of Ambiguous and Other Polysemous Utterances: Presented in Spoken Mode	12
	Comprehension of Ambiguous and Other Polysemous Utterances: Presented in Written Mode	15
Brown, Sandra	jt. auth., see Gary, Melvin	
Buckmaster, Annette	jt. auth., see Takala, Sauli	
Buxton, Amity P.	Dimensions for Looking at Children's Writings and Drawings in Daily Journals over Time	163
Cayer, Roger L.	Locus of Complexity in Written Language	166
Chin, Beverly Ann	The Chin Inventory on Content Area Reading Instruction	138
Clements, Paul	A Test of Staging Effects on Recall from Prose	69
Cohen, Arthur M.	Dichotomous Scale for Evaluating Expository Writing	168

Committee on Teaching and Its Evaluation in Composition of the Conference on College Composition and Communication	Evaluating Information in Composition	170
Cox, Carole A.	Film Preference Instrument	225
Cronnell, Bruce	jt. auth., see Humes, Ann	
Daly, John A.	Writing Apprehension Test	171
	jt. auth., see Witte, Stephen P.	
DeFord, Diane E.	The DeFord Theoretical Orientation to Reading Profile	141
Denburg, Susan Dalfen	The Interaction of Picture and Print in Reading Instruction	72
Devine, Thomas G.	Listening Skills Assessment	17
Duffy, Gerald	The Proposition Inventory	75
Dupuis, Mary M.	Criterion-Referenced Test for the Assessment of Reading and Writing Skills of Professional Educators	146
	1. Knowledge of Content Area Reading Skills; 2. Situations Survey: Teaching Reading in Content Areas; 3. Statements Survey: Teaching Reading in Content Areas	143
Emig, Janet	Emig-King Writing Attitude Scale for Students	173
	Emig-King Writing Attitude Scale for Teachers	148
Fagan, William T.	Comprehension Categories for Protocol Analysis	77
	The Syntactic Proposition for Protocol Analysis	80
	jt. auth., see Alparaque, Idrenne	
Faigley, Lester	jt. auth., see Witte, Stephen P.	
Fieker, Carolyn	jt. auth., see Humes, Ann	
Froese, Victor	The Quality, Direction, and Distance of Within Sentence Contextual Constraints	82
Gary, Melvin	Gary-Brown Writing Opinionnaire for College Instructors	150
Gentry, Larry	jt. auth., see Humes, Ann	
Gere, Anne Ruggles	jt. auth., see Schuessler, Brian F.	156
Gillis, Candida	A Survey of Methods and Materials	152
Goodman, Yetta M.	Techniques for Collecting Literacy Events from Young Children	19

Hanna, Peter	jt. auth., see Wilkinson, Andrew	
Hartnett, Carolyn	Analyzing Cohesive Ties	175
Hoffman, James V.	Feedback to Oral Reading Miscue Analysis System	85
Humes, Ann	Scoring Guides for Children's Writing	176
Indiana State Department of Public Instruction	Textbook Usage Inventory	87
King, Barbara	jt. auth., see Emig, Janet	
Koch, William R.	jt. auth., see Witte, Stephen P.	
Koziol, Stephen M., Jr.	Oral Composing. Communication Skills. Pennsylvania Comprehensive Reading/Communication Arts Plan Assessment Survey III	61
	Responding to Literature. Communication Skills. Pennsylvania Comprehensive Reading/Communication Arts Plan Assessment Survey I	29
Kumar, Krishna	Social Relationships in Children's Stories	32
Lawlor, Joseph	jt. auth., see Humes, Ann	
Lee, Joyce W.	jt. auth., see Dupuis, Mary M.	
Luban, Nina	jt. auth., see Reigstad, Tom	
Matsuhashi, Ann	jt. auth., see Reigstad, Tom	
Maull, Ethel M.	jt. auth., see Blake, Howard E.	
Martin, Anne V.	Extended-Cloze Reading Skills Test of Hierarchical and Spatial/Chronological Ordering Ability	88
McLoughlin, Jeffrey	jt. auth., see Dupuis, Mary M.	
Miller, Michael	jt. auth., see Daly, John A.	
Moore, David W.	Morpheme Facilitation Test	91
National Assessment of Educational Progress Staff and Consultants	Analyzing Characters in Literature	35
	Analyzing Themes in Poetry	39
	Applying Criteria to Evaluate Literature	42
	Cohesion Scoring Guides	180
	General Response to Literature	46
	Guidelines for Categorizing Mechanics and Grammatical Errors	184
	Guides for Evaluating Expressive Writing	185
	Guides for Evaluating Persuasive Writing	189
	Sentence Combining Scoring Guides	193
	Qualities of Good Literature	51

Neininger, Wendy	jt. auth., see Gillis, Candida	
Nicholson, Tom	An Assessment of the Effects of Different Error Types on the Understanding of Connected Discourse	94
Oklahoma State Department of Education	Curriculum Review Handbook: Language Arts	3
Oliver, Jo Ellen	A Tentative Criterion-Referenced Test to Measure Thinking Processes, Form A and B	97
Palmer, Barbara C.	jt. auth., see Smith, Edwin H.	
Pavlik, Robert A.	Mastery Assessment of Basic Reading Concepts	154
Pedersen, Elray L.	1. The Standard English as a Second Language Spoken English Test; 2. The Standard English as a Second Language Grammar Test; 3. The Standard English as a Second Language Vocabulary Test	21
	The Standard Test of Reading Effectiveness	100
Phillips-Riggs, Linda	Categories of Inferencing Strategies	101
Purves, Alan C.	jt. auth., see Takala, Sauli	
Quellmalz, Edys	Scale for Evaluating Expository Writing	196
	Scale for Evaluating Narrative Writing	198
Rabianski, Nancyanne	Scoring Writing with an Informative Aim	201
Reigstad, Tom	Writing Center Tutorial Record Form	203
Rose, Mike	Questionnaire for Identifying Writer's Block	204
Rosen, Lois	jt. auth., see Gillis, Candida	
Rubin, Rebecca B.	Communication Competency Assessment Instrument	63
Sacks, Renee K.	jt. auth., see Cayer, Roger L.	
Sadoski, Mark C.	Comprehension Process Score	103
Schuessler, Brian F.	Teacher Attitudes toward Composition/Instruction	156
Smith, Edwin H.	Smith/Palmer Figurative Language Interpretation Test	106
Snyder, Sandra L.	Rating Scale for the Assessment of the Speaking Skills of Teachers	158
	jt. auth., see Dupuis, Mary M.	
Stahl, Abraham	Analyzing the Structure of Children's Compositions	206
The State Education Department, New York	Degrees of Reading Power	108

Stice, Carole Kirchner	Test of Oral Contrastive Stress	111
Sticht, Thomas G.	Literacy Assessment Battery	25
Stone, David Edey	Test of Picture-Text Amalgams in Procedural Texts	112
Strange, Michael Caley	Orthographic Anomalies in a Silent Reading Task	114
Swan, Margaret	jt. auth., see Wilkinson, Andrew	
Swayne, Philip E.	Story Preference Inventory	55
Sword, Jeane	Criteria for Evaluating Picture Story Books	57
Takala, Sauli	Measures of a Writer's Choice	209
	Unipolar Scale for Evaluating Writing	212
Walmsley, Sean A.	Walmsley CVC Patterns Test	117
White, Jane	Taxonomy of Reading Behaviors	120
Wilkinson, Andrew	Models for Analysis of Writing	215
Williams, Connie K.	A Procedure to Evaluate Cognitive Requirements of Beginning Reading Materials	125
Witte, Stephen P.	Evaluating Course and Teacher Effectiveness	220
Young, Carlotta Joyner	jt. auth., see Dupuis, Mary M.	
Young, Petey	Discourse-Analysis Based, Written, Multiple-Choice Post-Test for Comprehension Assessment of Expository Prose	128
	An Instrument for Assessing Comprehension through the Ability to Recognize Verbatim Phrases from Previously Read Expository Prose	130

Category Index

Language Arts

Author	*Title*	*Page*
Oklahoma State Department of Education	Curriculum Review Handbook: Language Arts	3

Language and Language Development

Author	*Title*	*Page*
Black, J. K.	Interactional Competency Checklist	7
Blake, H. E. Maull, E. M.	Diagnosis of Language Competency Inventory	10
Brause, R. S.	Comprehension of Ambiguous and Other Polysemous Utterances: Presented in Spoken Mode	12
Brause, R. S.	Comprehension of Ambiguous and Other Polysemous Utterances: Presented in Written Mode	15
Devine, T. G.	Listening Skills Assessment	17
Goodman, Y. M.	Techniques for Collecting Literacy Events from Young Children	19
Pedersen, E. L.	The Standard English as a Second Language Spoken English Test	21
Pedersen, E. L.	The Standard English as a Second Language Grammar Test	21
Pedersen, E. L.	The Standard English as a Second Language Vocabulary Test	21

Literacy

Author	*Title*	*Page*
Sticht, T. G.	Literacy Assessment Battery	25

Literature

Author	*Title*	*Page*
Koziol, S. M., Jr.	Responding to Literature. Communication Skills. Pennsylvania Comprehensive Reading/Communication Arts Plan Assessment Survey I	29

Kumar, K.	Social Relationships in Children's Stories	32
National Assessment of Educational Progress Staff and Consultants	Analyzing Characters in Literature	35
National Assessment of Educational Progress Staff and Consultants	Analyzing Themes in Poetry	39
National Assessment of Educational Progress Staff and Consultants	Applying Criteria to Evaluate Literature	42
National Assessment of Educational Progress Staff and Consultants	General Response to Literature	46
National Assessment of Educational Progress Staff and Consultants	Qualities of Good Literature	51
Swayne, P. E.	Story Preference Inventory	55
Sword, J.	Criteria for Evaluating Picture Story Books	57

Oral Language

Author	*Title*	*Page*
Koziol, S. M., Jr.	Oral Composing. Communication Skills. Pennsylvania Comprehensive Reading/Communication Arts Plan Assessment Survey III	61
Rubin, R. B.	Communication Competency Assessment Instrument	63

Reading

Author	*Title*	*Page*
Alparaque, I. Fagan, W. T.	Assessment of Instructional Terms	67
Clements, P.	A Test of Staging Effects on Recall from Prose	69
Denburg, S. D.	The Interaction of Picture and Print in Reading Instruction	72
Duffy, G. Anderson, L.	The Proposition Inventory	75
Fagan, W. T.	Comprehension Categories for Protocol Analysis	77
Fagan, W. T.	The Syntactic Proposition for Protocol Analysis	80
Froese, V.	The Quality, Direction, and Distance of Within Sentence Contextual Constraints	82

Hoffman, J. V. Baker, C. J.	Feedback to Oral Reading Miscue Analysis System	85
Indiana State Department of Public Instruction	Textbook Usage Inventory	87
Martin, A. V.	Extended-Cloze Reading Skills Test of Hierarchical and Spatial/Chronological Ordering Ability	88
Moore, D. W.	Morpheme Facilitation Test	91
Nicholson, T.	An Assessment of the Effects of Different Error Types on the Understanding of Connected Discourse	94
Oliver, J. E.	A Tentative Criterion-Referenced Test to Measure Thinking Processes, Form A and B	97
Pedersen, E. L.	The Standard Test of Reading Effectiveness	100
Phillips-Riggs, L.	Categories of Inferencing Strategies	101
Sadoski, M. C.	Comprehension Process Score	103
Smith, E. H. Palmer, B. C.	Smith/Palmer Figurative Language Interpretation Test	106
State Education Department, New York	Degrees of Reading Power	108
Stice, C. K.	Test of Oral Contrastive Stress	111
Stone, D. E.	Test of Picture-Text Amalgams in Procedural Texts	112
Strange, M. C.	Orthographic Anomalies in a Silent Reading Task	114
Walmsley, S. A.	Walmsley CVC Patterns Test	117
White, J.	Taxonomy of Reading Behaviors	120
Williams, C. K.	A Procedure to Evaluate Cognitive Requirements of Beginning Reading Materials	125
Young, P.	Discourse-Analysis Based, Written, Multiple-Choice Post-Test for Comprehension Assessment of Expository Prose	128
Young, P.	An Instrument for Assessing Comprehension through the Ability to Recognize Verbatim Phrases from Previously Read Expository Prose	130

Teacher Knowledge/Attitudes

Author	*Title*	*Page*
Benterud, J. G.	Categories for Observing Language Arts Instruction	135

Chin, B. A.	The Chin Inventory on Content Area Reading Instruction	138
DeFord, D. E.	The DeFord Theoretical Orientation to Reading Profile	141
Dupuis, M. M. Askov, E. N. Lee, J. W. McLoughlin, J.	Knowledge of Content Area Reading Skills	143
Dupuis, M. M. Askov, E. N. Lee, J. W. Young, C. J.	Situations Survey: Teaching Reading in Content Areas	143
Dupuis, M. M. Askov, E. N. Lee, J. W. Young, C. J.	Statements Survey: Teaching Reading in Content Areas	143
Dupuis, M. M. Snyder, S. L.	Criterion-Referenced Test for the Assessment of Reading and Writing Skills of Professional Educators	146
Emig, J. King, B.	Emig-King Writing Attitude Scale for Teachers	148
Gary, M. Brown, S.	Gary-Brown Writing Opinionnaire for College Instructors	150
Gillis, C. Rosen, L. Neininger, W.	A Survey of Methods and Materials	152
Pavlik, R. A.	Mastery Assessment of Basic Reading Concepts	154
Schuessler, B. F. Gere, A. R. Abbott, R. D.	Teacher Attitudes toward Composition Instruction	156
Snyder, S. L.	Rating Scale for the Assessment of the Speaking Skills of Teachers	158

Writing

Author	*Title*	*Page*
Buxton, A. P.	Dimensions for Looking at Children's Writings and Drawings in Daily Journals over Time	163
Cayer, R. L. Sacks, R. K.	Locus of Complexity in Written Language	166
Cohen, A. M.	Dichotomous Scale for Evaluating Expository Writing	168

Conference on College Composition and Communication, Committee on Teaching and Its Evaluation in Composition	Evaluating Information in Composition	170
Daly, J. A. Miller, M.	Writing Apprehension Test	171
Emig, J. King, B.	Emig-King Writing Attitude Scale for Students	173
Hartnett, C.	Analyzing Cohesive Ties	175
Humes, A. Cronnell, B. Lawlor, J. Gentry, L. Fieker, C.	Scoring Guides for Children's Writing	176
National Assessment of Educational Progress Consultants and Staff	Cohesion Scoring Guides	180
National Assessment of Educational Progress Consultants and Staff	Guidelines for Categorizing Mechanics and Grammatical Errors	184
National Assessment of Educational Progress Consultants and Staff	Guides for Evaluating Expressive Writing	185
National Assessment of Educational Progress Consultants and Staff	Guides for Evaluating Persuasive Writing	189
National Assessment of Educational Progress Consultants and Staff	Sentence Combining Scoring Guides	193
Quellmalz, E.	Scale for Evaluating Expository Writing	196
Quellmalz, E. Center for the Study of Evaluation, University of California, Los Angeles	Scale for Evaluating Narrative Writing	198
Rabianski, N.	Scoring Writing with an Informative Aim	201
Reigstad, T. Matsuhashi, A. Luban, N.	Writing Center Tutorial Record Form	203
Rose, M.	Questionnaire for Identifying Writer's Block	204
Stahl, A.	Analyzing the Structure of Children's Compositions	206

Takala, S. Purves, A. C. Buckmaster, A.	Measures of a Writer's Choice	209
Takala, S. Purves, A. C. Buckmaster, A.	Unipolar Scale for Evaluating Writing	212
Wilkinson, A. Barnsley, G. Hanna, P. Swan, M.	Models for Analysis of Writing	215
Witte, S. P. Daly, J. A. Faigley, L. Koch, W. R.	Evaluating Course and Teacher Effectiveness	220

Miscellaneous

Author	*Title*	*Page*
Cox, C. A.	Film Preference Instrument	225